

1. Preface to BMCC Concepts of Biology 101, 2e
2. Introduction to Science

1. Introduction
2. The Science of Biology
3. Themes and Concepts of Biology

3. The Chemical Foundation of Life

1. Introduction
2. Atoms, Isotopes, Ions, and Molecules: The Building Blocks
3. Water
4. Carbon

4. Biological Macromolecules

1. Introduction
2. Synthesis of Biological Macromolecules
3. Carbohydrates
4. Lipids
5. Proteins
6. Nucleic Acids

5. The Cell

1. Cell Structure

1. Introduction
2. Studying Cells
3. Prokaryotic Cells
4. Eukaryotic Cells
5. The Endomembrane System and Proteins
6. The Cytoskeleton
7. Connections between Cells and Cellular Activities

2. Structure and Function of Plasma Membranes

1. Introduction

2. [Components and Structure](#)
3. [Passive Transport](#)
4. [Active Transport](#)
5. [Bulk Transport](#)

3. Cell Reproduction

1. [Introduction](#)
2. [Cell Division](#)
3. [The Cell Cycle](#)
4. [Control of the Cell Cycle](#)
5. [Cancer and the Cell Cycle](#)

6. Genetics

1. Meiosis and Sexual Reproduction

1. [Introduction](#)
2. [The Process of Meiosis](#)
3. [Sexual Reproduction](#)

2. Mendel's Experiments and Heredity

1. [Introduction](#)
2. [Mendel's Experiments and the Laws of Probability](#)
3. [Characteristics and Traits](#)
4. [Laws of Inheritance](#)

3. Modern Understandings of Inheritance

1. [Introduction](#)
2. [Chromosomal Theory and Genetic Linkage](#)
3. [Chromosomal Basis of Inherited Disorders](#)

4. DNA Structure and Function

1. [Introduction](#)
2. [Historical Basis of Modern Understanding](#)
3. [DNA Structure and Sequencing](#)
4. [Basics of DNA Replication](#)
5. [DNA Replication in Prokaryotes](#)

6. DNA Replication in Eukaryotes

7. DNA Repair

5. Genes and Proteins

1. Introduction

2. The Genetic Code

3. Prokaryotic Transcription

4. Eukaryotic Transcription

5. RNA Processing in Eukaryotes

6. Ribosomes and Protein Synthesis

6. Gene Expression

1. Introduction

2. Regulation of Gene Expression

3. Prokaryotic Gene Regulation

4. Eukaryotic Epigenetic Gene Regulation

5. Eukaryotic Transcription Gene Regulation

6. Eukaryotic Post-transcriptional Gene Regulation

7. Eukaryotic Translational and Post-translational
Gene Regulation

8. Cancer and Gene Regulation

7. Biotechnology and Genomics

1. Introduction

2. Biotechnology

3. Mapping Genomes

4. Whole-Genome Sequencing

5. Applying Genomics

6. Genomics and Proteomics

7. Evolutionary Processes

1. Evolution and the Origin of Species

1. Introduction

2. Understanding Evolution

3. Formation of New Species

4. Reconnection and Speciation Rates

2. The Evolution of Populations

1. [Introduction](#)
 2. [Population Evolution](#)
 3. [Population Genetics](#)
 4. [Adaptive Evolution](#)
-
8. [The Periodic Table of Elements](#)
 9. [Measurements and the Metric System](#)

Preface to BMCC Concepts of Biology 101, 2e

Welcome to this collection of materials from *Biology 2e* (2nd edition), an OpenStax resource. This textbook was written to increase student access to high-quality learning materials, maintaining highest standards of academic rigor at little to no cost.

About OpenStax

OpenStax is a nonprofit based at Rice University, and it's our mission to improve student access to education. Our first openly licensed college textbook was published in 2012, and our library has since scaled to over 25 books for college and AP® courses used by hundreds of thousands of students.

OpenStax Tutor, our low-cost personalized learning tool, is being used in college courses throughout the country. Through our partnerships with philanthropic foundations and our alliance with other educational resource organizations, OpenStax is breaking down the most common barriers to learning and empowering students and instructors to succeed.

About OpenStax resources

Customization

Biology 2e is licensed under a Creative Commons Attribution 4.0 International (CC BY) license, which means that you can distribute, remix, and build upon the content, as long as you provide attribution to OpenStax and its content contributors.

Because our books are openly licensed, you are free to use the entire book or pick and choose the sections that are most relevant to the needs of your course. Feel free to remix the content by assigning your students certain chapters and sections in your syllabus, in the order that you prefer. You can even provide a direct link in your syllabus to the sections in the web view of your book.

Instructors also have the option of creating a customized version of their OpenStax book. The custom version can be made available to students in low-cost print or digital form through their campus bookstore. Visit the Instructor Resources section of your book page on OpenStax.org for more information.

Errata

All OpenStax textbooks undergo a rigorous review process. However, like any professional-grade textbook, errors sometimes occur. Since our books are web based, we can make updates periodically

when deemed pedagogically necessary. If you have a correction to suggest, submit it through the link on your book page on OpenStax.org. Subject matter experts review all errata suggestions. OpenStax is committed to remaining transparent about all updates, so you will also find a list of past errata changes on your book page on OpenStax.org.

Format

You can access this textbook for free in web view or PDF through OpenStax.org, and for a low cost in print.

About the authors

Second edition authors and reviewers

Senior Contributing Authors

Mary Ann Clark, Texas Wesleyan University

Jung Choi, Georgia Institute of Technology

Matthew Douglas, Grand Rapids Community College

Reviewers

Kathleen Berlyn, Baltimore City Community College

Bridgett Brinton, Armstrong State University

Jennifer Chase, Northwest Nazarene University

Amy Hoffman, Grayson County College

Olga Kopp, Utah Valley University

Jennifer Larson, Capital University
Jason Locklin, Austin Community College
Hongmei Ma, American University
Melissa Masse, Tulsa Community College
Shannon McDermott, Central Virginia Community College
Bryan Monesson-Olson, University of Massachusetts Amherst
Amber Reece, California State University Fresno
Monique Reed, Texas A&M University
Jeffrey Roberts, American River College
Matthew Smith, North Dakota State University
Dawn Wankowski, Cardinal Stritch University

First edition authors and reviewers

Senior Contributing Authors

Yael Avissar (Cell Biology), Rhode Island College
Jung Choi (Genetics), Georgia Institute of Technology
Jean DeSaix (Evolution), University of North Carolina at Chapel Hill
Vladimir Jurukovski (Animal Physiology), Suffolk County Community College
Robert Wise (Plant Biology), University of Wisconsin, Oshkosh
Connie Rye (General Content Lead), East Mississippi Community College

Contributing Authors and Reviewers

Julie Adams, Aurora University

Summer Allen, Brown University
James Bader, Case Western Reserve University
David Bailey, St. Norbert College
Mark Belk, Brigham Young University
Nancy Boury, Iowa State University
Lisa Bonneau, Metropolitan Community College –
Blue River
Graciela Brelles-Marino, California State University
Pomona
Mark Browning, Purdue University
Sue Chaplin, University of St. Thomas
George Cline, Jacksonville State University
Deb Cook, Georgia Gwinnett College
Diane Day, Clayton State University
Frank Dirrigl, The University of Texas Pan American
Waneene Dorsey, Grambling State University
Nick Downey, University of Wisconsin La Crosse
Rick Duhrkopf, Baylor University
Kristy Duran, Adams State University
Stan Eisen, Christian Brothers University
Brent Ewers, University of Wyoming
Myriam Feldman, Lake Washington Institute of
Technology
Michael Fine, Virginia Commonwealth University
Linda Flora, Delaware County Community College
Thomas Freeland, Walsh University
David Gris , Texas A & M University – Corpus
Christi
Andrea Hazard, SUNY Cortland
Michael Hedrick, University of North Texas
Linda Hensel, Mercer University

Mark Kopeny, University of Virginia
Norman Johnson, University of Massachusetts
Amherst
Grace Lasker, Lake Washington Institute of
Technology; Walden University
Sandy Latourelle, SUNY Plattsburgh
Theo Light, Shippensburg University
Clark Lindgren, Grinnell College
James Malcolm, University of Redlands
Mark Meade, Jacksonville State University
Richard Merritt, Houston Community College
James Mickle, North Carolina State University
Jasleen Mishra, Houston Community College
Dudley Moon, Albany College of Pharmacy and
Health Sciences
Shobhana Natarajan, Brookhaven College
Jonas Okeagu, Fayetteville State University
Diana Oliveras, University of Colorado Boulder
John Peters, College of Charleston
Joel Piperberg, Millersville University
Johanna Porter-Kelley, Winston-Salem State
University
Robyn Puffenbarger, Bridgewater College
Dennis Revie, California Lutheran University
Ann Rushing, Baylor University
Sangha Saha, City College of Chicago
Edward Saiff, Ramapo College of New Jersey
Brian Shmaefsky, Lone Star College System
Robert Sizemore, Alcorn State University
Marc Smith, Sinclair Community College
Frederick Spiegel, University of Arkansas

Frederick Sproull, La Roche College
Bob Sullivan, Marist College
Mark Sutherland, Hendrix College
Toure Thompson, Alabama A&M University
Scott Thomson, University of Wisconsin – Parkside
Allison van de Meene, University of Melbourne
Mary White, Southeastern Louisiana University
Steven Wilt, Bellarmine University
James Wise, Hampton University
Renna Wolfe
Virginia Young, Mercer University
Leslie Zeman, University of Washington
Daniel Zurek, Pittsburg State University
Shobhana Natarajan, Alcon Laboratories, Inc.

Introduction

class = "introduction" This NASA image is a composite of several satellite-based views of Earth. To make the whole-Earth image, NASA scientists combine observations of different parts of the planet. (credit: NASA/GSFC/NOAA/USGS)



Viewed from space, Earth offers no clues about the diversity of life forms that reside there. Scientists believe that the first forms of life on Earth were microorganisms that existed for billions of years in the ocean before plants and animals appeared. The mammals, birds, and flowers so familiar to us are all relatively recent, originating 130 to 250 million years ago. The earliest representatives of the genus *Homo*, to which we belong, have inhabited this planet for only the last 2.5 million years, and only in the last 300,000 years have humans started looking like we do today.

The Science of Biology

By the end of this section, you will be able to do the following:

- Identify the shared characteristics of the natural sciences
- Summarize the steps of the scientific method
- Compare inductive reasoning with deductive reasoning
- Describe the goals of basic science and applied science

Formerly called blue-green algae, these (a) cyanobacteria, magnified 300x under a light microscope, are some of Earth's oldest life forms. These (b) stromatolites along the shores of Lake Thetis in Western Australia are ancient structures formed by layering cyanobacteria in shallow waters. (credit a: modification of work by NASA; credit b: modification of work by Ruth Ellison; scale-bar data from Matt Russell)



(a)

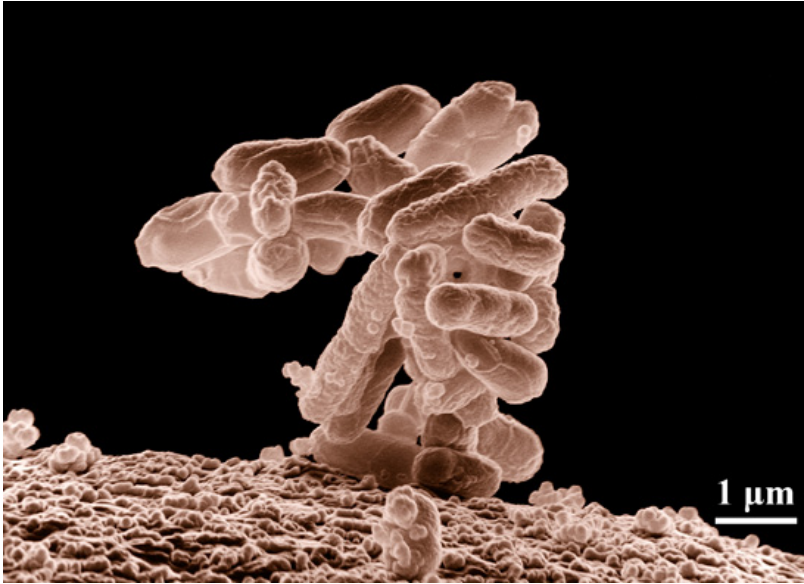


(b)

What is biology? In simple terms, **biology** is the study of living organisms and their interactions with one another and their environments. This is a very

broad definition because the scope of biology is vast. Biologists may study anything from the microscopic or submicroscopic view of a cell to ecosystems and the whole living planet ([\[link\]](#)). Listening to the daily news, you will quickly realize how many aspects of biology we discuss every day. For example, recent news topics include *Escherichia coli* ([\[link\]](#)) outbreaks in spinach and *Salmonella* contamination in peanut butter. Other subjects include efforts toward finding a cure for AIDS, Alzheimer's disease, and cancer. On a global scale, many researchers are committed to finding ways to protect the planet, solve environmental issues, and reduce the effects of climate change. All of these diverse endeavors are related to different facets of the discipline of biology.

Escherichia coli (*E. coli*) bacteria, in this scanning electron micrograph, are normal residents of our digestive tracts that aid in absorbing vitamin K and other nutrients. However, virulent strains are sometimes responsible for disease outbreaks. (credit: Eric Erbe, digital colorization by Christopher Pooley, both of USDA, ARS, EMU)



The diversity of scientific fields includes astronomy, biology, computer science, geology, logic, physics, chemistry, mathematics, and many other fields.
(credit: “Image Editor”/Flickr)

The Process of Science

Biology is a science, but what exactly is science? What does the study of biology share with other scientific disciplines? We can define **science** (from the Latin *scientia*, meaning “knowledge”) as knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method. It becomes clear from this definition that applying scientific method plays a major role in science. The **scientific method** is a method of research with defined steps that include experiments and careful observation.

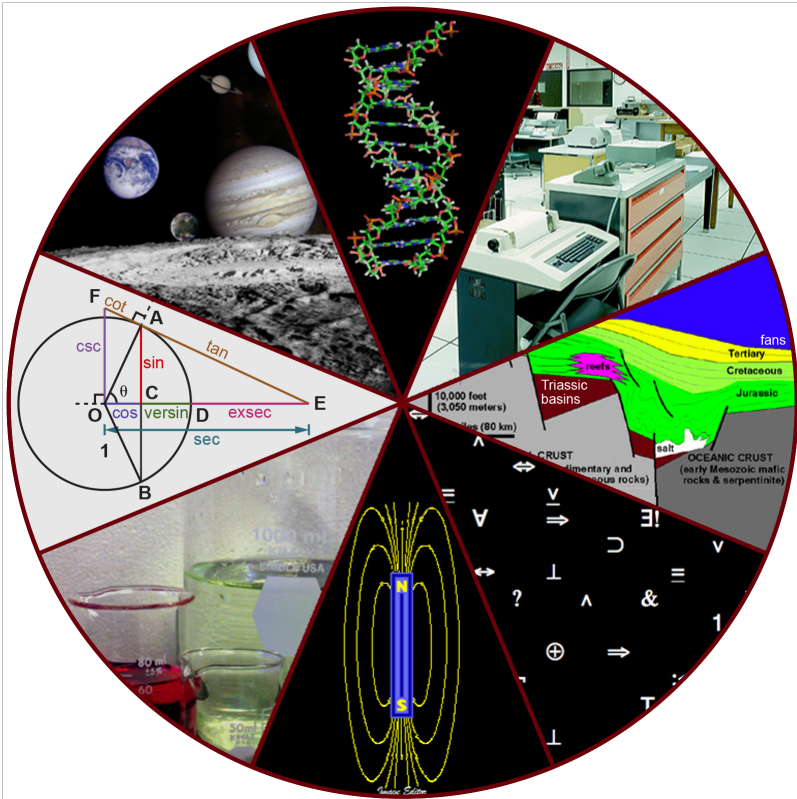
We will examine scientific method steps in detail later, but one of the most important aspects of this method is the testing of hypotheses by means of repeatable experiments. A **hypothesis** is a suggested explanation for an event, which one can test.

Although using the scientific method is inherent to science, it is inadequate in determining what science is. This is because it is relatively easy to apply the scientific method to disciplines such as physics and chemistry, but when it comes to disciplines like archaeology, psychology, and geology, the scientific method becomes less applicable as repeating experiments becomes more difficult.

These areas of study are still sciences, however. Consider archaeology—even though one cannot perform repeatable experiments, hypotheses may still be supported. For instance, an archaeologist can hypothesize that an ancient culture existed based on finding a piece of pottery. He or she could make further hypotheses about various characteristics of this culture, which could be correct or false through continued support or contradictions from other findings. A hypothesis may become a verified theory. A **theory** is a tested and confirmed explanation for observations or phenomena. Therefore, we may be better off to define science as fields of study that attempt to comprehend the nature of the universe.

Natural Sciences

What would you expect to see in a museum of natural sciences? Frogs? Plants? Dinosaur skeletons? Exhibits about how the brain functions? A planetarium? Gems and minerals? Maybe all of the above? Science includes such diverse fields as astronomy, biology, computer sciences, geology, logic, physics, chemistry, and mathematics ([\[link\]](#)). However, scientists consider those fields of science related to the physical world and its phenomena and processes **natural sciences**. Thus, a museum of natural sciences might contain any of the items listed above.



There is no complete agreement when it comes to defining what the natural sciences include, however. For some experts, the natural sciences are astronomy, biology, chemistry, earth science, and physics. Other scholars choose to divide natural sciences into **life sciences**, which study living things and include biology, and **physical sciences**, which study nonliving matter and include astronomy, geology, physics, and chemistry. Some disciplines such as biophysics and biochemistry build on both life and physical sciences and are interdisciplinary. Some refer to natural sciences as “hard science” because they rely on the use of quantitative data. Social sciences that study society and human behavior are more likely to use qualitative assessments to drive investigations and findings.

Not surprisingly, the natural science of biology has many branches or subdisciplines. Cell biologists study cell structure and function, while biologists who study anatomy investigate the structure of an entire organism. Those biologists studying physiology, however, focus on the internal functioning of an organism. Some areas of biology focus on only particular types of living things. For example, botanists explore plants, while zoologists specialize in animals.

Scientific Reasoning

One thing is common to all forms of science: an ultimate goal “to know.” Curiosity and inquiry are the driving forces for the development of science. Scientists seek to understand the world and the way it operates. To do this, they use two methods of logical thinking: inductive reasoning and deductive reasoning.

Inductive reasoning is a form of logical thinking that uses related observations to arrive at a general conclusion. This type of reasoning is common in descriptive science. A life scientist such as a biologist makes observations and records them. These data can be qualitative or quantitative, and one can supplement the raw data with drawings, pictures, photos, or videos. From many observations, the scientist can infer conclusions (inductions) based on evidence. Inductive reasoning involves formulating generalizations inferred from careful observation and analyzing a large amount of data. Brain studies provide an example. In this type of research, scientists observe many live brains while people are engaged in a specific activity, such as viewing images of food. The scientist then predicts the part of the brain that “lights up” during this activity to be the part controlling the response to the selected stimulus, in this case, images of food. Excess absorption of radioactive sugar derivatives by active areas of the brain causes the various areas to “light up”. Scientists use a scanner to observe the resultant increase in radioactivity. Then, researchers

can stimulate that part of the brain to see if similar responses result.

Deductive reasoning or deduction is the type of logic used in hypothesis-based science. In deductive reason, the pattern of thinking moves in the opposite direction as compared to inductive reasoning. **Deductive reasoning** is a form of logical thinking that uses a general principle or law to forecast specific results. From those general principles, a scientist can extrapolate and predict the specific results that would be valid as long as the general principles are valid. Studies in climate change can illustrate this type of reasoning. For example, scientists may predict that if the climate becomes warmer in a particular region, then the distribution of plants and animals should change.

Both types of logical thinking are related to the two main pathways of scientific study: descriptive science and hypothesis-based science. **Descriptive (or discovery) science**, which is usually inductive, aims to observe, explore, and discover, while **hypothesis-based science**, which is usually deductive, begins with a specific question or problem and a potential answer or solution that one can test. The boundary between these two forms of study is often blurred, and most scientific endeavors combine both approaches. The fuzzy boundary becomes apparent when thinking about how easily observation can lead to specific questions. For

example, a gentleman in the 1940s observed that the burr seeds that stuck to his clothes and his dog's fur had a tiny hook structure. On closer inspection, he discovered that the burrs' gripping device was more reliable than a zipper. He eventually experimented to find the best material that acted similar, and produced the hook-and-loop fastener popularly known today as Velcro. Descriptive science and hypothesis-based science are in continuous dialogue.

Historians credit Sir Francis Bacon (1561–1626) as the first to define the scientific method. (credit: Paul van Somer)

The Scientific Method

Biologists study the living world by posing questions about it and seeking science-based responses. Known as scientific method, this approach is common to other sciences as well. The scientific method was used even in ancient times, but England's Sir Francis Bacon (1561–1626) first documented it ([\[link\]](#)). He set up inductive methods for scientific inquiry. The scientific method is not used only by biologists; researchers from almost all fields of study can apply it as a logical, rational problem-solving method.



The scientific process typically starts with an observation (often a problem to solve) that leads to a question. Let's think about a simple problem that starts with an observation and apply the scientific method to solve the problem. One Monday morning, a student arrives at class and quickly discovers that the classroom is too warm. That is an observation that also describes a problem: the classroom is too warm. The student then asks a question: "Why is the classroom so warm?"

Proposing a Hypothesis

Recall that a hypothesis is a suggested explanation that one can test. To solve a problem, one can propose several hypotheses. For example, one hypothesis might be, “The classroom is warm because no one turned on the air conditioning.” However, there could be other responses to the question, and therefore one may propose other hypotheses. A second hypothesis might be, “The classroom is warm because there is a power failure, and so the air conditioning doesn’t work.”

Once one has selected a hypothesis, the student can make a prediction. A prediction is similar to a hypothesis but it typically has the format “If . . . then” For example, the prediction for the first hypothesis might be, “*If* the student turns on the air conditioning, *then* the classroom will no longer be too warm.”

Testing a Hypothesis

A valid hypothesis must be testable. It should also be **falsifiable**, meaning that experimental results can disprove it. Importantly, science does not claim to “prove” anything because scientific understandings are always subject to modification with further information. This step—openness to disproving ideas—is what distinguishes sciences from non-sciences. The presence of the supernatural,

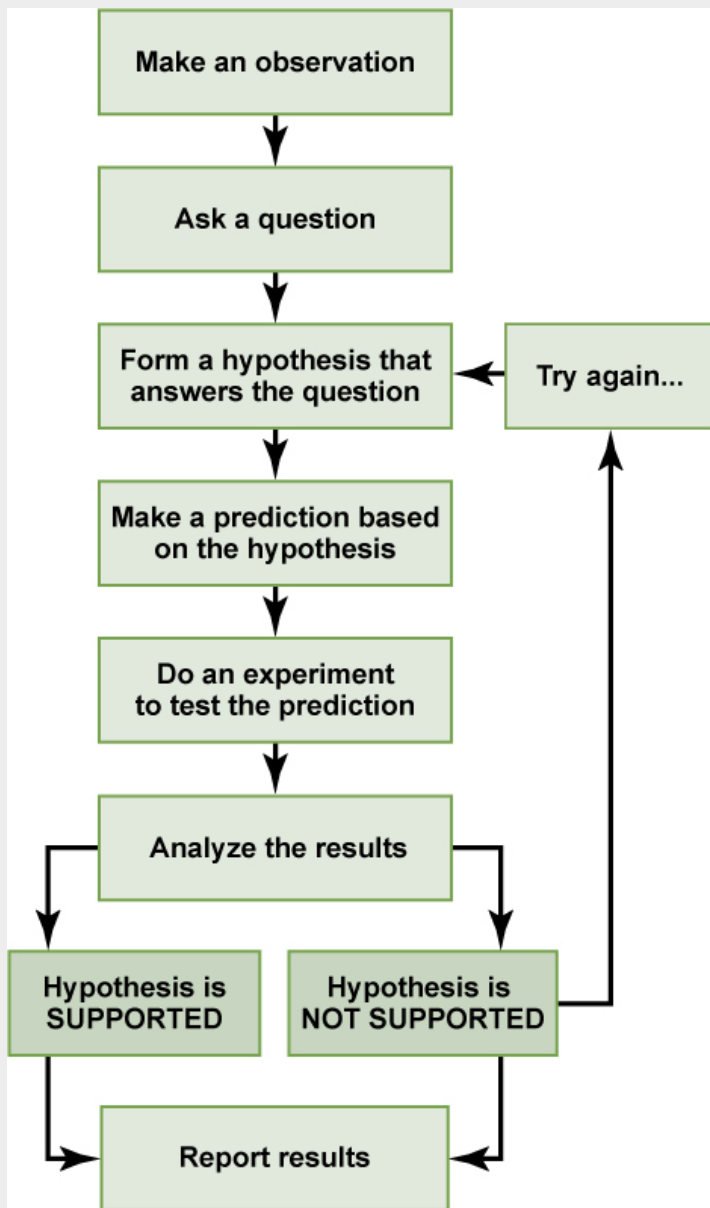
for instance, is neither testable nor falsifiable. To test a hypothesis, a researcher will conduct one or more experiments designed to eliminate one or more of the hypotheses. Each experiment will have one or more variables and one or more controls. A **variable** is any part of the experiment that can vary or change during the experiment. The **control group** contains every feature of the experimental group except it is not given the manipulation that the researcher hypothesizes. Therefore, if the experimental group's results differ from the control group, the difference must be due to the hypothesized manipulation, rather than some outside factor. Look for the variables and controls in the examples that follow. To test the first hypothesis, the student would find out if the air conditioning is on. If the air conditioning is turned on but does not work, there should be another reason, and the student should reject this hypothesis. To test the second hypothesis, the student could check if the lights in the classroom are functional. If so, there is no power failure and the student should reject this hypothesis. The students should test each hypothesis by carrying out appropriate experiments. Be aware that rejecting one hypothesis does not determine whether or not one can accept the other hypotheses. It simply eliminates one hypothesis that is not valid ([\[link\]](#)). Using the scientific method, the student rejects the hypotheses that are inconsistent with experimental data.

While this “warm classroom” example is based on observational results, other hypotheses and experiments might have clearer controls. For instance, a student might attend class on Monday and realize she had difficulty concentrating on the lecture. One observation to explain this occurrence might be, “When I eat breakfast before class, I am better able to pay attention.” The student could then design an experiment with a control to test this hypothesis.

In hypothesis-based science, researchers predict specific results from a general premise. We call this type of reasoning deductive reasoning: deduction proceeds from the general to the particular. However, the reverse of the process is also possible: sometimes, scientists reach a general conclusion from a number of specific observations. We call this type of reasoning inductive reasoning, and it proceeds from the particular to the general. Researchers often use inductive and deductive reasoning in tandem to advance scientific knowledge ([\[link\]](#)).

Visual Connection

The scientific method consists of a series of well-defined steps. If a hypothesis is not supported by experimental data, one can propose a new hypothesis.



In the example below, the scientific method is used to solve an everyday problem. Match the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items).

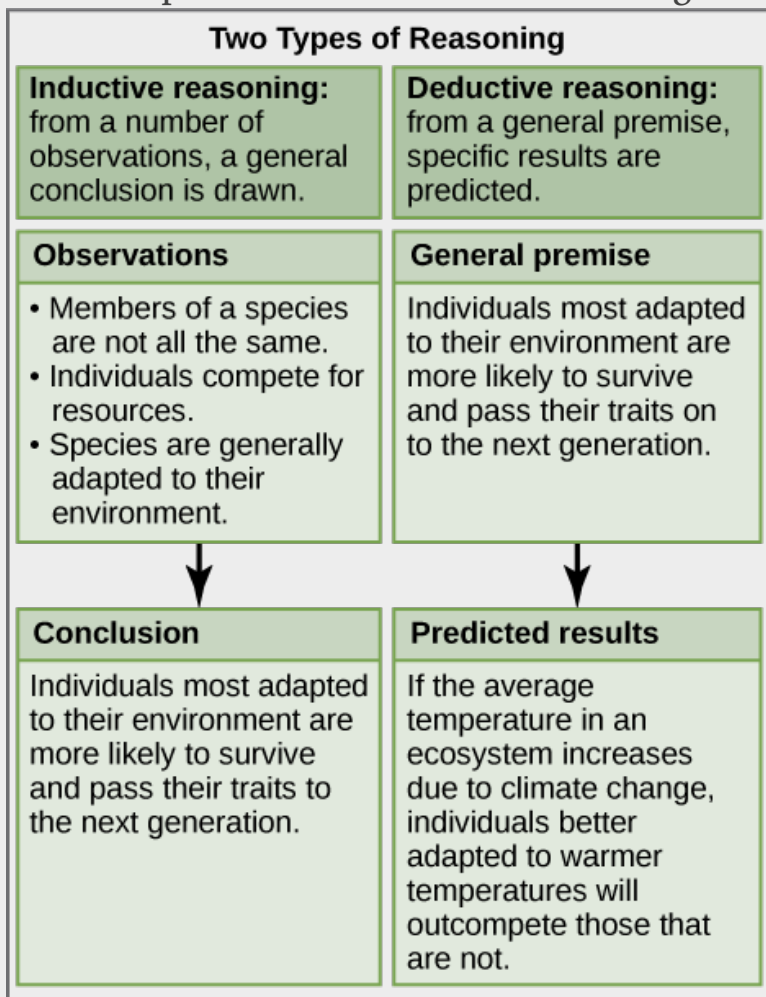
Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation	a. There is something wrong with the electrical outlet.
2. Question	b. If something is wrong with the outlet, my coffeemaker also won't work when plugged into it.
3. Hypothesis (answer)	c. My toaster doesn't toast my bread.
4. Prediction	d. I plug my coffee maker into the outlet.
5. Experiment	e. My coffeemaker works.
6. Result	f. Why doesn't my toaster work?

Visual Connection

Scientists use two types of reasoning, inductive and deductive reasoning, to advance scientific knowledge. As is the case in this example, the

conclusion from inductive reasoning can often become the premise for deductive reasoning.



Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.

2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA, are distributed evenly between the daughter cells during cell division. Therefore, each daughter cell will have the same chromosome set as the mother cell.
4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

The scientific method may seem too rigid and structured. It is important to keep in mind that, although scientists often follow this sequence, there is flexibility. Sometimes an experiment leads to conclusions that favor a change in approach. Often, an experiment brings entirely new scientific questions to the puzzle. Many times, science does not operate in a linear fashion. Instead, scientists continually draw inferences and make generalizations, finding patterns as their research proceeds. Scientific reasoning is more complex than the scientific method alone suggests. Notice, too, that we can apply the scientific method to solving problems that aren't necessarily scientific in nature. After Hurricane Irma struck the Caribbean and

Florida in 2017, thousands of baby squirrels like this one were thrown from their nests. Thanks to applied science, scientists knew how to rehabilitate the squirrel. (credit: audreyjm529, Flickr) The Human Genome Project was a 13-year collaborative effort among researchers working in several different science fields. Researchers completed the project, which sequenced the entire human genome, in 2003. (credit: the U.S. Department of Energy Genome Programs (<http://genomics.energy.gov>))

Two Types of Science: Basic Science and Applied Science

The scientific community has been debating for the last few decades about the value of different types of science. Is it valuable to pursue science for the sake of simply gaining knowledge, or does scientific knowledge only have worth if we can apply it to solving a specific problem or to bettering our lives? This question focuses on the differences between two types of science: basic science and applied science.

Basic science or “pure” science seeks to expand knowledge regardless of the short-term application of that knowledge. It is not focused on developing a product or a service of immediate public or commercial value. The immediate goal of basic science is knowledge for knowledge’s sake, although this does not mean that, in the end, it may not result

in a practical application.

In contrast, **applied science** or “technology,” aims to use science to solve real-world problems, making it possible, for example, to improve a crop yield, find a cure for a particular disease, or save animals threatened by a natural disaster ([\[link\]](#)). In applied science, the problem is usually defined for the researcher.



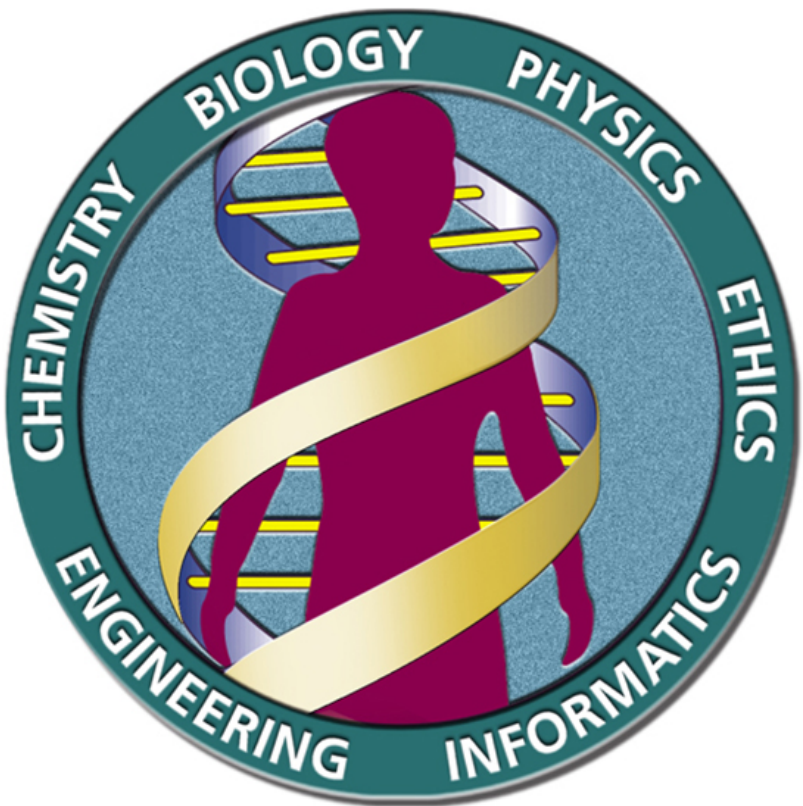
Some individuals may perceive applied science as “useful” and basic science as “useless.” A question these people might pose to a scientist advocating knowledge acquisition would be, “What for?” However, a careful look at the history of science reveals that basic knowledge has resulted in many remarkable applications of great value. Many

scientists think that a basic understanding of science is necessary before researchers develop an application therefore, applied science relies on the results that researchers generate through basic science. Other scientists think that it is time to move on from basic science in order to find solutions to actual problems. Both approaches are valid. It is true that there are problems that demand immediate attention; however, scientists would find few solutions without the help of the wide knowledge foundation that basic science generates.

One example of how basic and applied science can work together to solve practical problems occurred after the discovery of DNA structure led to an understanding of the molecular mechanisms governing DNA replication. DNA strands, unique in every human, are in our cells, where they provide the instructions necessary for life. When DNA replicates, it produces new copies of itself, shortly before a cell divides. Understanding DNA replication mechanisms enabled scientists to develop laboratory techniques that researchers now use to identify genetic diseases, pinpoint individuals who were at a crime scene, and determine paternity. Without basic science, it is unlikely that applied science would exist.

Another example of the link between basic and applied research is the Human Genome Project, a study in which researchers analyzed and mapped

each human chromosome to determine the precise sequence of DNA subunits and each gene's exact location. (The gene is the basic unit of heredity. An individual's complete collection of genes is his or her genome.) Researchers have studied other less complex organisms as part of this project in order to gain a better understanding of human chromosomes. The Human Genome Project ([\[link\]](#)) relied on basic research with simple organisms and, later, with the human genome. An important end goal eventually became using the data for applied research, seeking cures and early diagnoses for genetically related diseases.



While scientists usually carefully plan research efforts in both basic science and applied science, note that some discoveries are made by **serendipity**, that is, by means of a fortunate accident or a lucky surprise. Scottish biologist Alexander Fleming discovered penicillin when he accidentally left a petri dish of *Staphylococcus* bacteria open. An unwanted mold grew on the dish, killing the bacteria. Fleming's curiosity to investigate the reason behind the bacterial death, followed by his experiments, led to the discovery of the antibiotic penicillin, which is produced by the fungus *Penicillium*. Even in the highly organized world of science, luck—when combined with an observant, curious mind—can lead to unexpected breakthroughs.

Reporting Scientific Work

Whether scientific research is basic science or applied science, scientists must share their findings in order for other researchers to expand and build upon their discoveries. Collaboration with other scientists—when planning, conducting, and analyzing results—are all important for scientific research. For this reason, important aspects of a scientist's work are communicating with peers and disseminating results to peers. Scientists can share results by presenting them at a scientific meeting or conference, but this approach can reach only the

select few who are present. Instead, most scientists present their results in peer-reviewed manuscripts that are published in scientific journals. **Peer-reviewed manuscripts** are scientific papers that a scientist's colleagues or peers review. These colleagues are qualified individuals, often experts in the same research area, who judge whether or not the scientist's work is suitable for publication. The process of peer review helps to ensure that the research in a scientific paper or grant proposal is original, significant, logical, and thorough. Grant proposals, which are requests for research funding, are also subject to peer review. Scientists publish their work so other scientists can reproduce their experiments under similar or different conditions to expand on the findings. The experimental results must be consistent with the findings of other scientists.

A scientific paper is very different from creative writing. Although creativity is required to design experiments, there are fixed guidelines when it comes to presenting scientific results. First, scientific writing must be brief, concise, and accurate. A scientific paper needs to be succinct but detailed enough to allow peers to reproduce the experiments.

The scientific paper consists of several specific sections—introduction, materials and methods, results, and discussion. This structure is sometimes called the “IMRaD” format. There are usually

acknowledgment and reference sections as well as an **abstract** (a concise summary) at the beginning of the paper. There might be additional sections depending on the type of paper and the journal where it will be published. For example, some review papers require an outline.

The **introduction** starts with brief, but broad, background information about what is known in the field. A good introduction also gives the rationale of the work. It justifies the work carried out and also briefly mentions the end of the paper, where the researcher will present the hypothesis or research question driving the research. The introduction refers to the published scientific work of others and therefore requires citations following the style of the journal. Using the work or ideas of others without proper citation is **plagiarism**.

The **materials and methods** section includes a complete and accurate description of the substances the researchers use, and the method and techniques they use to gather data. The description should be thorough enough to allow another researcher to repeat the experiment and obtain similar results, but it does not have to be verbose. This section will also include information on how the researchers made measurements and the types of calculations and statistical analyses they used to examine raw data. Although the materials and methods section gives an accurate description of the experiments, it does

not discuss them.

Some journals require a results section followed by a discussion section, but it is more common to combine both. If the journal does not allow combining both sections, the **results** section simply narrates the findings without any further interpretation. The researchers present results with tables or graphs, but they do not present duplicate information. In the **discussion** section, the researchers will interpret the results, describe how variables may be related, and attempt to explain the observations. It is indispensable to conduct an extensive literature search to put the results in the context of previously published scientific research. Therefore, researchers include proper citations in this section as well.

Finally, the **conclusion** section summarizes the importance of the experimental findings. While the scientific paper almost certainly answers one or more scientific questions that the researchers stated, any good research should lead to more questions. Therefore, a well-done scientific paper allows the researchers and others to continue and expand on the findings.

Review articles do not follow the IMRAD format because they do not present original scientific findings, or primary literature. Instead, they summarize and comment on findings that were

published as primary literature and typically include extensive reference sections.

Section Summary

Biology is the science that studies living organisms and their interactions with one another and their environments. Science attempts to describe and understand the nature of the universe in whole or in part by rational means. Science has many fields. Those fields related to the physical world and its phenomena are natural sciences.

Science can be basic or applied. The main goal of basic science is to expand knowledge without any expectation of short-term practical application of that knowledge. The primary goal of applied research, however, is to solve practical problems.

Science uses two types of logical reasoning. Inductive reasoning uses particular results to produce general scientific principles. Deductive reasoning is a form of logical thinking that predicts results by applying general principles. The common thread throughout scientific research is using the scientific method, a step-based process that consists of making observations, defining a problem, posing hypotheses, testing these hypotheses, and drawing one or more conclusions. The testing uses proper controls. Scientists present their results in peer-

reviewed scientific papers published in scientific journals. A scientific research paper consists of several well-defined sections: introduction, materials and methods, results, and, finally, a concluding discussion. Review papers summarize the conducted research in a particular field over a period of time.

Visual Connection Questions

[\[link\]](#) In the example below, the scientific method is used to solve an everyday problem. Match the scientific method steps (numbered items) with the process of solving the everyday problem (lettered items). Based on the results of the experiment, is the hypothesis correct? If it is incorrect, propose some alternative hypotheses.

1. Observation	a. There is something wrong with the electrical outlet.
2. Question	b. If something is wrong with the outlet, my coffeemaker also won't work when

	plugged into it.
3. Hypothesis (answer)	c. My toaster doesn't toast my bread.
4. Prediction	d. I plug my coffee maker into the outlet.
5. Experiment	e. My coffeemaker works.
6. Result	f. Why doesn't my toaster work?

[\[link\]](#) 1: C; 2: F; 3: A; 4: B; 5: D; 6: E. The original hypothesis is incorrect, as the coffeemaker works when plugged into the outlet. Alternative hypotheses include that the toaster might be broken or that the toaster wasn't turned on.

[\[link\]](#) Decide if each of the following is an example of inductive or deductive reasoning.

1. All flying birds and insects have wings. Birds and insects flap their wings as they move through the air. Therefore, wings enable flight.
2. Insects generally survive mild winters better than harsh ones. Therefore, insect pests will become more problematic if global temperatures increase.
3. Chromosomes, the carriers of DNA,

separate into daughter cells during cell division. Therefore, each daughter cell will have the same chromosome set as the mother cell.

4. Animals as diverse as humans, insects, and wolves all exhibit social behavior. Therefore, social behavior must have an evolutionary advantage.

[\[link\]](#) 1: inductive; 2: deductive; 3: deductive; 4: inductive.

Review Questions

The first forms of life on Earth were _____.

1. plants
2. microorganisms
3. birds
4. dinosaurs

B

A suggested and testable explanation for an event is called a _____.

1. hypothesis
 2. variable
 3. theory
 4. control
-

A

Which of the following sciences is not considered a natural science?

1. biology
 2. astronomy
 3. physics
 4. computer science
-

D

The type of logical thinking that uses related observations to arrive at a general conclusion is called _____.

1. deductive reasoning
 2. the scientific method
 3. hypothesis-based science
 4. inductive reasoning
-

D

The process of _____ helps to ensure that a scientist's research is original, significant, logical, and thorough.

1. publication
2. public speaking
3. peer review
4. the scientific method

C

A person notices that her houseplants that are regularly exposed to music seem to grow more quickly than those in rooms with no music. As a result, she determines that plants grow better when exposed to music. This example most closely resembles which type of reasoning?

1. inductive reasoning
2. deductive reasoning
3. neither, because no hypothesis was made
4. both inductive and deductive reasoning

A

Critical Thinking Questions

Although the scientific method is used by most of the sciences, it can also be applied to everyday situations. Think about a problem that you may have at home, at school, or with your car, and apply the scientific method to solve it.

Answers will vary, but should apply the steps of the scientific method. One possibility could be a car which doesn't start. The hypothesis could be that the car doesn't start because the battery is dead. The experiment would be to change the battery or to charge the battery and then check whether the car starts or not. If it starts, the problem was due to the battery, and the hypothesis is accepted.

Give an example of how applied science has had a direct effect on your daily life.

Answers will vary. One example of how applied science has had a direct effect on daily life is the presence of vaccines. Vaccines to prevent diseases such as polio, measles, tetanus, and even influenza affect daily life by contributing to individual and societal health.

Name two topics that are likely to be studied by biologists, and two areas of scientific study that

would fall outside the realm of biology.

Answers will vary. Topics that fall inside the area of biological study include how diseases affect human bodies, how pollution impacts a species' habitat, and how plants respond to their environments. Topics that fall outside of biology (the “study of life”) include how metamorphic rock is formed and how planetary orbits function.

Thinking about the topic of cancer, write a basic science question and an applied science question that a researcher interested in this topic might ask.

Answers will vary. Basic science: What evolutionary purpose might cancer serve?
Applied science: What strategies might be found to prevent cancer from reproducing at the cellular level?

Glossary

abstract

opening section of a scientific paper that summarizes the research and conclusions

applied science

form of science that aims to solve real-world problems

basic science

science that seeks to expand knowledge and understanding regardless of the short-term application of that knowledge

biology

the study of living organisms and their interactions with one another and their environments

conclusion

section of a scientific paper that summarizes the importance of the experimental findings

control

part of an experiment that does not change during the experiment

deductive reasoning

form of logical thinking that uses a general inclusive statement to forecast specific results

descriptive science

(also, discovery science) form of science that aims to observe, explore, and investigate

discussion

section of a scientific paper in which the

author interprets experimental results, describes how variables may be related, and attempts to explain the phenomenon in question

falsifiable

able to be disproven by experimental results

hypothesis

suggested explanation for an observation, which one can test

hypothesis-based science

form of science that begins with a specific question and potential testable answers

inductive reasoning

form of logical thinking that uses related observations to arrive at a general conclusion

introduction

opening section of a scientific paper, which provides background information about what was known in the field prior to the research reported in the paper

life science

field of science, such as biology, that studies living things

materials and methods

section of a scientific paper that includes a

complete description of the substances, methods, and techniques that the researchers used to gather data

natural science

field of science that is related to the physical world and its phenomena and processes

peer-reviewed manuscript

scientific paper that a scientist's colleagues review who are experts in the field of study

physical science

field of science, such as geology, astronomy, physics, and chemistry, that studies nonliving matter

plagiarism

using other people's work or ideas without proper citation, creating the false impression that those are the author's original ideas

results

section of a scientific paper in which the author narrates the experimental findings and presents relevant figures, pictures, diagrams, graphs, and tables, without any further interpretation

review article

paper that summarizes and comments on findings that were published as primary

literature

science

knowledge that covers general truths or the operation of general laws, especially when acquired and tested by the scientific method

scientific method

method of research with defined steps that include observation, formulation of a hypothesis, testing, and confirming or falsifying the hypothesis

serendipity

fortunate accident or a lucky surprise

theory

tested and confirmed explanation for observations or phenomena

variable

part of an experiment that the experimenter can vary or change

Themes and Concepts of Biology

By the end of this section, you will be able to do the following:

- Identify and describe the properties of life
- Describe the levels of organization among living things
- Recognize and interpret a phylogenetic tree
- List examples of different subdisciplines in biology

Biology is the science that studies life, but what exactly is life? This may sound like a silly question with an obvious response, but it is not always easy to define life. For example, a branch of biology called virology studies viruses, which exhibit some of the characteristics of living entities but lack others. Although viruses can attack living organisms, cause diseases, and even reproduce, they do not meet the criteria that biologists use to define life. Consequently, virologists are not biologists, strictly speaking. Similarly, some biologists study the early molecular evolution that gave rise to life. Since the events that preceded life are not biological events, these scientists are also excluded from biology in the strict sense of the term.

From its earliest beginnings, biology has wrestled with three questions: What are the shared properties that make something “alive”? Once we know something is alive, how do we find meaningful

levels of organization in its structure? Finally, when faced with the remarkable diversity of life, how do we organize the different kinds of organisms so that we can better understand them? As scientists discover new organisms every day, biologists continue to seek answers to these and other questions.

A toad represents a highly organized structure consisting of cells, tissues, organs, and organ systems. (credit: “Ivengo”/Wikimedia Commons) The leaves of this sensitive plant (*Mimosa pudica*) will instantly droop and fold when touched. After a few minutes, the plant returns to normal. (credit: Alex Lomas) Although no two look alike, these kittens have inherited genes from both parents and share many of the same characteristics. (credit: Rocky Mountain Feline Rescue) Polar bears (*Ursus maritimus*) and other mammals living in ice-covered regions maintain their body temperature by generating heat and reducing heat loss through thick fur and a dense layer of fat under their skin. (credit: “longhorndave”/Flickr) The California condor (*Gymnogyps californianus*) uses chemical energy derived from food to power flight. California condors are an endangered species. This bird has a wing tag that helps biologists identify the individual. (credit: Pacific Southwest Region U.S. Fish and Wildlife Service)

Properties of Life

All living organisms share several key characteristics or functions: order, sensitivity or response to the environment, reproduction, adaptation, growth and development, regulation, homeostasis, energy processing, and evolution. When viewed together, these nine characteristics serve to define life.

Order



Organisms are highly organized, coordinated structures that consist of one or more cells. Even very simple, single-celled organisms are remarkably complex: inside each cell, atoms comprise molecules. These in turn comprise cell organelles and other cellular inclusions. In multicellular organisms ([\[link\]](#)), similar cells form tissues. Tissues, in turn, collaborate to create organs (body structures with a distinct function). Organs work

together to form organ systems.

Sensitivity or Response to Stimuli



Organisms respond to diverse stimuli. For example, plants can bend toward a source of light, climb on fences and walls, or respond to touch ([\[link\]](#)). Even tiny bacteria can move toward or away from chemicals (a process called *chemotaxis*) or light (*phototaxis*). Movement toward a stimulus is a positive response, while movement away from a stimulus is a negative response.

Link to Learning

Watch [this video](#) to see how plants respond to a

stimulus—from opening to light, to wrapping a tendril around a branch, to capturing prey.

Reproduction

Single-celled organisms reproduce by first duplicating their DNA, and then dividing it equally as the cell prepares to divide to form two new cells. Multicellular organisms often produce specialized reproductive germline, gamete, oocyte, and sperm cells. After fertilization (the fusion of an oocyte and a sperm cell), a new individual develops. When reproduction occurs, DNA containing genes are passed along to an organism's offspring. These genes ensure that the offspring will belong to the same species and will have similar characteristics, such as size and shape.

Growth and Development

Organisms grow and develop as a result of genes providing specific instructions that will direct cellular growth and development. This ensures that a species' young ([\[link\]](#)) will grow up to exhibit many of the same characteristics as its parents.



Regulation

Even the smallest organisms are complex and require multiple regulatory mechanisms to coordinate internal functions, respond to stimuli, and cope with environmental stresses. Two examples of internal functions regulated in an organism are nutrient transport and blood flow. Organs (groups of tissues working together) perform specific functions, such as carrying oxygen throughout the body, removing wastes, delivering nutrients to every cell, and cooling the body.

Homeostasis



In order to function properly, cells require appropriate conditions such as proper temperature, pH, and appropriate concentration of diverse chemicals. These conditions may, however, change from one moment to the next. Organisms are able to maintain internal conditions within a narrow range almost constantly, despite environmental changes, through **homeostasis** (literally, “steady state”). For example, an organism needs to regulate body temperature through the thermoregulation process. Organisms that live in cold climates, such as the polar bear ([\[link\]](#)), have body structures that help them withstand low temperatures and conserve body heat. Structures that aid in this type of insulation include fur, feathers, blubber, and fat. In hot climates, organisms have methods (such as perspiration in humans or panting in dogs) that help them to shed excess body heat.

Energy Processing

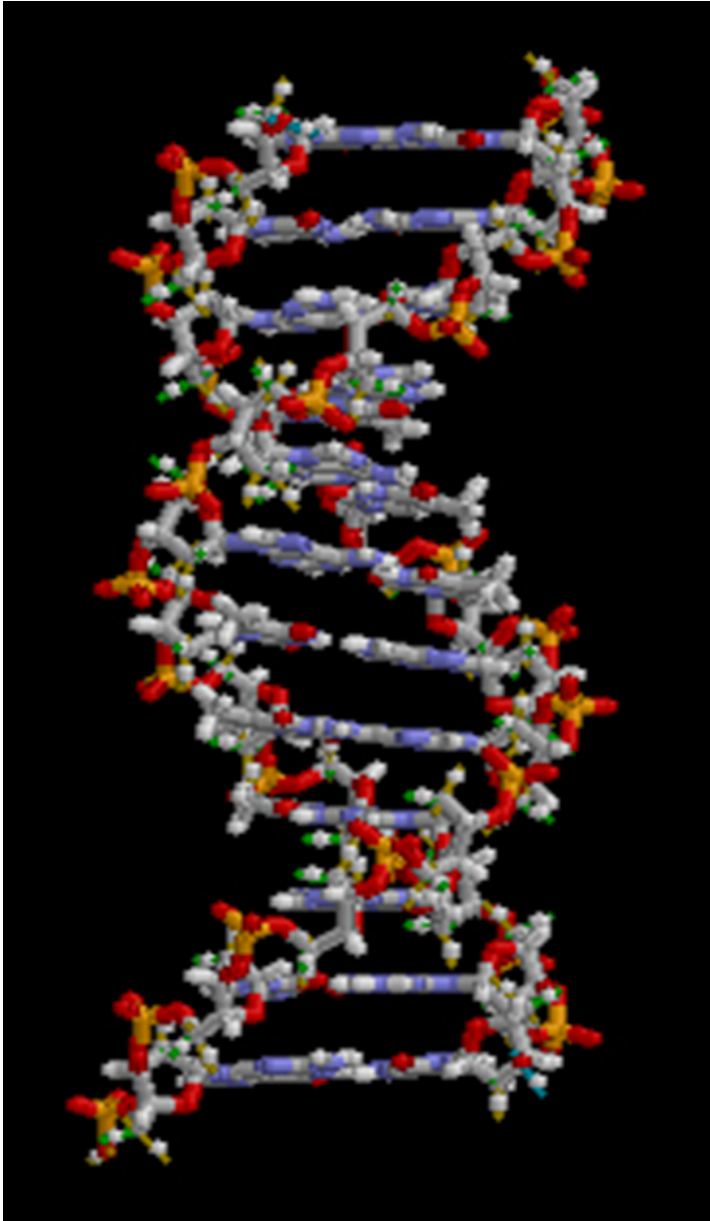


All organisms use a source of energy for their metabolic activities. Some organisms capture energy from the sun and convert it into chemical energy in food. Others use chemical energy in molecules they take in as food ([\[link\]](#)).

All molecules, including this DNA molecule, are comprised of atoms. (credit: “brian0918”/Wikimedia Commons)

Levels of Organization of Living Things

Living things are highly organized and structured, following a hierarchy that we can examine on a scale from small to large. The **atom** is the smallest and most fundamental unit of matter. It consists of a nucleus surrounded by electrons. Atoms form molecules. A **molecule** is a chemical structure consisting of at least two atoms held together by one or more chemical bonds. Many molecules that are biologically important are **macromolecules**, large molecules that are typically formed by polymerization (a polymer is a large molecule that is made by combining smaller units called monomers, which are simpler than macromolecules). An example of a macromolecule is deoxyribonucleic acid (DNA) ([\[link\]](#)), which contains the instructions for the structure and functioning of all living organisms.



Link to Learning

Watch [this video](#) that animates the three-dimensional structure of the DNA molecule in [\[link\]](#).

Some cells contain aggregates of macromolecules surrounded by membranes. We call these **organelles**. Organelles are small structures that exist within cells. Examples of organelles include mitochondria and chloroplasts, which carry out indispensable functions: mitochondria produce energy to power the cell, while chloroplasts enable green plants to utilize the energy in sunlight to make sugars. All living things are made of cells. The **cell** itself is the smallest fundamental unit of structure and function in living organisms. (This requirement is why scientists do not consider viruses living: they are not made of cells. To make new viruses, they have to invade and hijack the reproductive mechanism of a living cell. Only then can they obtain the materials they need to reproduce.) Some organisms consist of a single cell and others are multicellular. Scientists classify cells as prokaryotic or eukaryotic. **Prokaryotes** are single-celled or colonial organisms that do not have membrane-bound nuclei. In contrast, the cells of **eukaryotes** do have membrane-bound organelles and a membrane-bound nucleus.

In larger organisms, cells combine to make **tissues**,

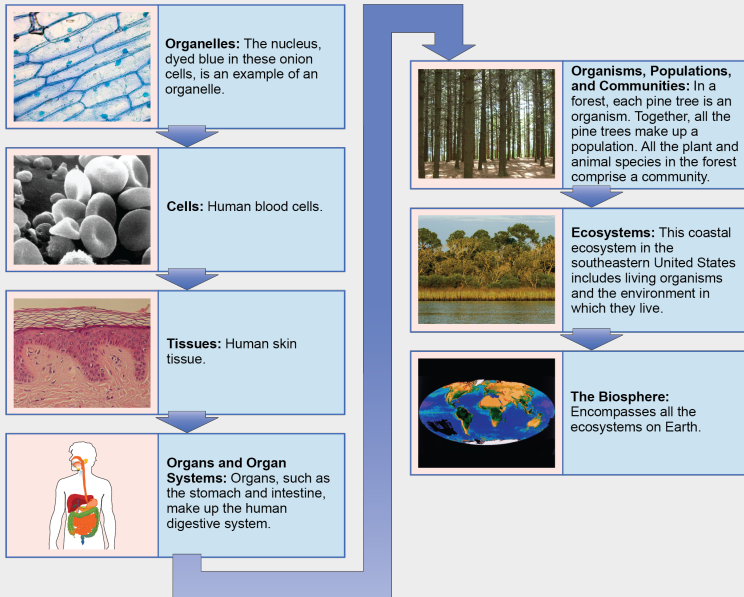
which are groups of similar cells carrying out similar or related functions. **Organs** are collections of tissues grouped together performing a common function. Organs are present not only in animals but also in plants. An **organ system** is a higher level of organization that consists of functionally related organs. Mammals have many organ systems. For instance, the circulatory system transports blood through the body and to and from the lungs. It includes organs such as the heart and blood vessels. **Organisms** are individual living entities. For example, each tree in a forest is an organism. Single-celled prokaryotes and single-celled eukaryotes are also organisms, which biologists typically call microorganisms.

Biologists collectively call all the individuals of a species living within a specific area a **population**. For example, a forest may include many pine trees, which represent the population of pine trees in this forest. Different populations may live in the same specific area. For example, the forest with the pine trees includes populations of flowering plants, insects, and microbial populations. A **community** is the sum of populations inhabiting a particular area. For instance, all of the trees, flowers, insects, and other populations in a forest form the forest's community. The forest itself is an ecosystem. An **ecosystem** consists of all the living things in a particular area together with the abiotic, nonliving parts of that environment such as nitrogen in the

soil or rain water. At the highest level of organization ([\[link\]](#)), the **biosphere** is the collection of all ecosystems, and it represents the zones of life on Earth. It includes land, water, and even the atmosphere to a certain extent.

Visual Connection

shows the biological levels of organization of living things. From a single organelle to the entire biosphere, living organisms are parts of a highly structured hierarchy. (credit “organelles”: modification of work by Umberto Salvagnin; credit “cells”: modification of work by Bruce Wetzels, Harry Schaefer/ National Cancer Institute; credit “tissues”: modification of work by Kilbad; Fama Clamosa; Mikael Häggström; credit “organs”: modification of work by Mariana Ruiz Villareal; credit “organisms”: modification of work by "Crystal"/Flickr; credit “ecosystems”: modification of work by US Fish and Wildlife Service Headquarters; credit “biosphere”: modification of work by NASA)



Which of the following statements is false?

1. Tissues exist within organs which exist within organ systems.
2. Communities exist within populations which exist within ecosystems.
3. Organelles exist within cells which exist within tissues.
4. Communities exist within ecosystems which exist in the biosphere.

Microbiologist Carl Woese constructed this phylogenetic tree using data that he obtained from sequencing ribosomal RNA genes. The tree shows the separation of living organisms into three domains: Bacteria, Archaea, and Eukarya. Bacteria

and Archaea are prokaryotes, single-celled organisms lacking intracellular organelles. (credit: Eric Gaba; NASA Astrobiology Institute)

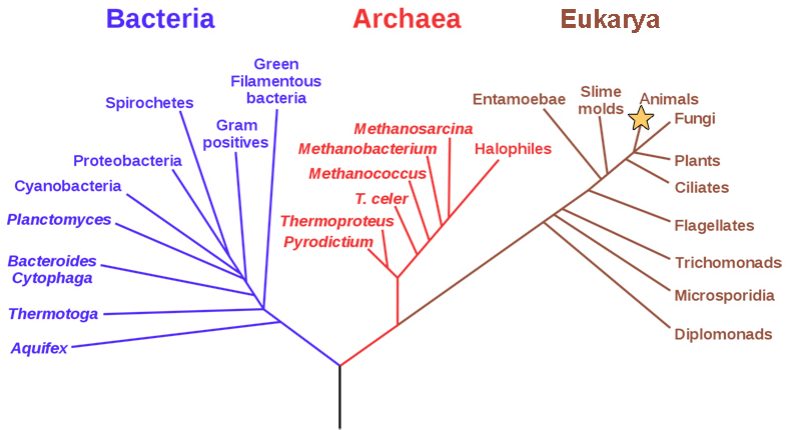
The Diversity of Life

The fact that biology, as a science, has such a broad scope has to do with the tremendous diversity of life on earth. The source of this diversity is **evolution**, the process of gradual change in a population or species over time. Evolutionary biologists study the evolution of living things in everything from the microscopic world to ecosystems.

A phylogenetic tree ([\[link\]](#)) can summarize the evolution of various life forms on Earth. It is a diagram showing the evolutionary relationships among biological species based on similarities and differences in genetic or physical traits or both. Nodes and branches comprise a phylogenetic tree. The internal nodes represent ancestors and are points in evolution when, based on scientific evidence, researchers believe an ancestor has diverged to form two new species. The length of each branch is proportional to the time elapsed since the split.

Phylogenetic Tree of Life

★ = You are here



Evolution Connection

Carl Woese and the Phylogenetic Tree

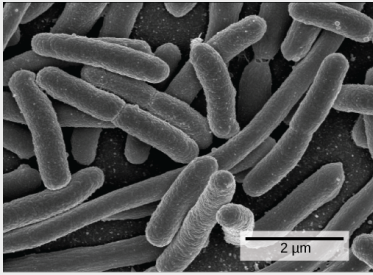
In the past, biologists grouped living organisms into five kingdoms: animals, plants, fungi, protists, and bacteria. They based the organizational scheme mainly on physical features, as opposed to physiology, biochemistry, or molecular biology, all of which modern systematics use. American microbiologist Carl Woese's pioneering work in the early 1970s has shown, however, that life on Earth has evolved along three lineages, now called domains—Bacteria, Archaea, and Eukarya. The first two are prokaryotic cells with microbes that lack membrane-enclosed nuclei and organelles. The third domain contains the eukaryotes and includes unicellular microorganisms (protists), together with

the three remaining kingdoms (fungi, plants, and animals). Woese defined Archaea as a new domain, and this resulted in a new taxonomic tree ([\[link\]](#)). Many organisms belonging to the Archaea domain live under extreme conditions and are called extremophiles. To construct his tree, Woese used genetic relationships rather than similarities based on morphology (shape).

Woese constructed his tree from universally distributed comparative gene sequencing that are present in every organism, and conserved (meaning that these genes have remained essentially unchanged throughout evolution). Woese's approach was revolutionary because comparing physical features are insufficient to differentiate between the prokaryotes that appear fairly similar in spite of their tremendous biochemical diversity and genetic variability ([\[link\]](#)). Comparing homologous DNA and RNA sequences provided Woese with a sensitive device that revealed the extensive variability of prokaryotes, and which justified separating the prokaryotes into two domains: bacteria and archaea.

These images represent different domains. The (a) bacteria in this micrograph belong to Domain Bacteria, while the (b) extremophiles (not visible) living in this hot vent belong to Domain Archaea. Both the (c) sunflower and (d) lion are part of Domain Eukarya. (credit a: modification of work by Drew March; credit b: modification of work by Steve Jurvetson; credit c: modification of work by

Michael Arrighi; credit d: modification of work by Leszek Leszcynski)



(a)



(b)



(c)



(d)

Researchers work on excavating dinosaur fossils at a site in Castellón, Spain. (credit: Mario Modesto)

Branches of Biological Study

The scope of biology is broad and therefore contains many branches and subdisciplines. Biologists may pursue one of those subdisciplines and work in a

more focused field. For instance, **molecular biology** and **biochemistry** study biological processes at the molecular and chemical level, including interactions among molecules such as DNA, RNA, and proteins, as well as the way they are regulated.

Microbiology, the study of microorganisms, is the study of the structure and function of single-celled organisms. It is quite a broad branch itself, and depending on the subject of study, there are also microbial physiologists, ecologists, and geneticists, among others.

Career Connection

Forensic Scientist

Forensic science is the application of science to answer questions related to the law. Biologists as well as chemists and biochemists can be forensic scientists. Forensic scientists provide scientific evidence for use in courts, and their job involves examining trace materials associated with crimes. Interest in forensic science has increased in the last few years, possibly because of popular television shows that feature forensic scientists on the job. Also, developing molecular techniques and establishing DNA databases have expanded the types of work that forensic scientists can do. Their job activities are primarily related to crimes against people such as murder, rape, and assault. Their work involves analyzing samples such as hair,

blood, and other body fluids and also processing DNA ([link](#)) found in many different environments and materials. Forensic scientists also analyze other biological evidence left at crime scenes, such as insect larvae or pollen grains. Students who want to pursue careers in forensic science will most likely have to take chemistry and biology courses as well as some intensive math courses.

This forensic scientist works in a DNA extraction room at the U.S. Army Criminal Investigation Laboratory at Fort Gillem, GA. (credit: United States Army CID Command Public Affairs)



Another field of biological study, **neurobiology**, studies the biology of the nervous system, and although it is a branch of biology, it is also an interdisciplinary field of study known as

neuroscience. Because of its interdisciplinary nature, this subdiscipline studies different nervous system functions using molecular, cellular, developmental, medical, and computational approaches.



Paleontology, another branch of biology, uses fossils to study life's history ([\[link\]](#)). **Zoology** and **botany** are the study of animals and plants, respectively. Biologists can also specialize as biotechnologists, ecologists, or physiologists, to name just a few areas. This is just a small sample of the many fields that biologists can pursue.

Biology is the culmination of the achievements of the natural sciences from their inception to today. Excitingly, it is the cradle of emerging sciences, such as the biology of brain activity, genetic engineering of custom organisms, and the biology of evolution

that uses the laboratory tools of molecular biology to retrace the earliest stages of life on Earth. A scan of news headlines—whether reporting on immunizations, a newly discovered species, sports doping, or a genetically-modified food—demonstrates the way biology is active in and important to our everyday world.

Section Summary

Biology is the science of life. All living organisms share several key properties such as order, sensitivity or response to stimuli, reproduction, growth and development, regulation, homeostasis, and energy processing. Living things are highly organized parts of a hierarchy that includes atoms, molecules, organelles, cells, tissues, organs, and organ systems. In turn, biologists group organisms as populations, communities, ecosystems, and the biosphere. The great diversity of life today evolved from less-diverse ancestral organisms over billions of years. We can use a phylogenetic tree to show evolutionary relationships among organisms.

Biology is very broad and includes many branches and subdisciplines. Examples include molecular biology, microbiology, neurobiology, zoology, and botany, among others.

Visual Connection Questions

[\[link\]](#) Which of the following statements is false?

1. Tissues exist within organs which exist within organ systems.
 2. Communities exist within populations which exist within ecosystems.
 3. Organelles exist within cells which exist within tissues.
 4. Communities exist within ecosystems which exist in the biosphere.
-

[\[link\]](#) Communities exist within populations which exist within ecosystems.

Review Questions

The smallest unit of biological structure that meets the functional requirements of “living” is the _____.

1. organ
2. organelle
3. cell

4. macromolecule

C

Viruses are not considered living because they _____.

1. are not made of cells
 2. lack cell nuclei
 3. do not contain DNA or RNA
 4. cannot reproduce
-

A

The presence of a membrane-enclosed nucleus is a characteristic of _____.

1. prokaryotic cells
 2. eukaryotic cells
 3. living organisms
 4. bacteria
-

B

A group of individuals of the same species living in the same area is called a(n) _____.

1. family
 2. community
 3. population
 4. ecosystem
-

C

Which of the following sequences represents the hierarchy of biological organization from the most inclusive to the least complex level?

1. organelle, tissue, biosphere, ecosystem, population
 2. organ, organism, tissue, organelle, molecule
 3. organism, community, biosphere, molecule, tissue, organ
 4. biosphere, ecosystem, community, population, organism
-

D

Where in a phylogenetic tree would you expect to find the organism that had evolved most recently?

1. at the base
2. within the branches

3. at the nodes
4. at the branch tips

D

Critical Thinking Questions

Select two items that biologists agree are necessary in order to consider an organism “alive.” For each, give an example of a nonliving object that otherwise fits the definition of “alive.”

Answers will vary. Layers of sedimentary rock have order but are not alive. Technology is capable of regulation but is not, of itself, alive.

Consider the levels of organization of the biological world, and place each of these items in order from smallest level of organization to most encompassing: skin cell, elephant, water molecule, planet Earth, tropical rainforest, hydrogen atom, wolf pack, liver.

Smallest level of organization to largest:

hydrogen atom, water molecule, skin cell, liver, elephant, wolf pack, tropical rainforest, planet Earth

You go for a long walk on a hot day. Give an example of a way in which homeostasis keeps your body healthy.

During your walk, you may begin to perspire, which cools your body and helps your body to maintain a constant internal temperature. You might also become thirsty and pause long enough for a cool drink, which will help to restore the water lost during perspiration.

Using examples, explain how biology can be studied from a microscopic approach to a global approach.

Researchers can approach biology from the smallest to the largest, and everything in between. For instance, an ecologist may study a population of individuals, the population's community, the community's ecosystem, and the ecosystem's part in the biosphere. When studying an individual organism, a biologist could examine the cell and its organelles, the tissues that the cells make up, the organs and

their respective organ systems, and the sum total—the organism itself.

Glossary

atom

smallest and most fundamental unit of matter

biochemistry

study of the chemistry of biological organisms

biosphere

collection of all the ecosystems on Earth

botany

study of plants

cell

smallest fundamental unit of structure and function in living things

community

set of populations inhabiting a particular area

ecosystem

all the living things in a particular area together with the abiotic, nonliving parts of that environment

eukaryote

organism with cells that have nuclei and

membrane-bound organelles

evolution

the process of gradual change in a population or species over time

homeostasis

ability of an organism to maintain constant internal conditions

macromolecule

large molecule, typically formed by the joining of smaller molecules

microbiology

study of the structure and function of microorganisms

molecule

chemical structure consisting of at least two atoms held together by one or more chemical bonds

molecular biology

study of biological processes and their regulation at the molecular level, including interactions among molecules such as DNA, RNA, and proteins

neurobiology

study of the biology of the nervous system

organ

collection of related tissues grouped together
performing a common function

organ system

level of organization that consists of
functionally related interacting organs

organelle

small structures that exist within cells and
carry out cellular functions

organism

individual living entity

paleontology

study of life's history by means of fossils

phylogenetic tree

diagram showing the evolutionary
relationships among various biological species
based on similarities and differences in
genetic or physical traits or both; in essence, a
hypothesis concerning evolutionary
connections

population

all of the individuals of a species living within
a specific area

prokaryote

single-celled organism that lacks organelles
and does not have nuclei surrounded by a

nuclear membrane

tissue

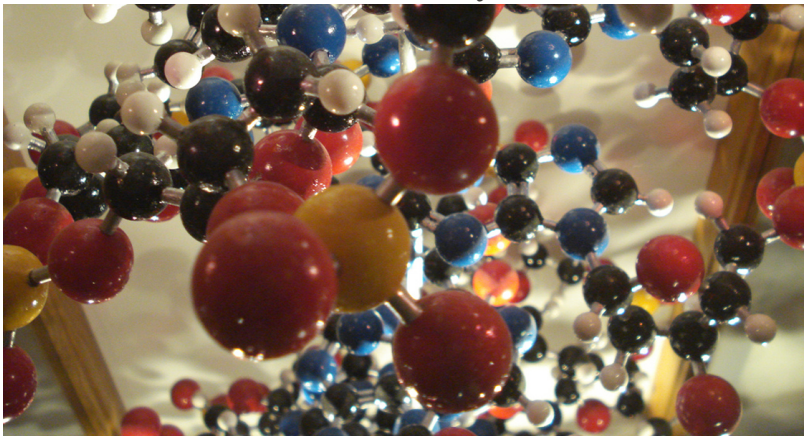
group of similar cells carrying out related functions

zoology

study of animals

Introduction

class = "introduction" Atoms are the building blocks of molecules in the universe—air, soil, water, rocks . . . and also the cells of all living organisms. In this model of an organic molecule, the atoms of carbon (black), hydrogen (white), nitrogen (blue), oxygen (red), and sulfur (yellow) are in proportional atomic size. The silver rods indicate chemical bonds. (credit: modification of work by Christian Guthier)



Elements in various combinations comprise all matter, including living things. Some of the most abundant elements in living organisms include carbon, hydrogen, nitrogen, oxygen, sulfur, and phosphorus. These form the nucleic acids, proteins, carbohydrates, and lipids that are the fundamental components of living matter. Biologists must understand these important building blocks and the unique structures of the atoms that comprise molecules, allowing for cells, tissues, organ systems, and entire organisms to form.

All biological processes follow the laws of physics and chemistry, so in order to understand how biological systems work, it is important to understand the underlying physics and chemistry. For example, the flow of blood within the circulatory system follows the laws of physics that regulate the modes of fluid flow. The breakdown of the large, complex molecules of food into smaller molecules—and the conversion of these to release energy to be stored in adenosine triphosphate (ATP)—is a series of chemical reactions that follow chemical laws. The properties of water and the formation of hydrogen bonds are key to understanding living processes. Recognizing the properties of acids and bases is important, for example, to our understanding of the digestive process. Therefore, the fundamentals of physics and chemistry are important for gaining insight into biological processes.

Atoms, Isotopes, Ions, and Molecules: The Building Blocks

By the end of this section, you will be able to do the following:

- Define matter and elements
- Describe the interrelationship between protons, neutrons, and electrons
- Compare the ways in which electrons can be donated or shared between atoms
- Explain the ways in which naturally occurring elements combine to create molecules, cells, tissues, organ systems, and organisms

At its most fundamental level, life is made up of matter. **Matter** is any substance that occupies space and has mass. **Elements** are unique forms of matter with specific chemical and physical properties that cannot break down into smaller substances by ordinary chemical reactions. There are 118 elements, but only 98 occur naturally. The remaining elements are unstable and require scientists to synthesize them in laboratories.

Each element is designated by its chemical symbol, which is a single capital letter or, when the first letter is already “taken” by another element, a combination of two letters. Some elements follow the English term for the element, such as C for carbon and Ca for calcium. Other elements’ chemical symbols derive from their Latin names. For

example, the symbol for sodium is Na, referring to *natrium*, the Latin word for sodium.

The four elements common to all living organisms are oxygen (O), carbon (C), hydrogen (H), and nitrogen (N). In the nonliving world, elements are found in different proportions, and some elements common to living organisms are relatively rare on the earth as a whole, as [\[link\]](#) shows. For example, the atmosphere is rich in nitrogen and oxygen but contains little carbon and hydrogen, while the earth's crust, although it contains oxygen and a small amount of hydrogen, has little nitrogen and carbon. In spite of their differences in abundance, all elements and the chemical reactions between them obey the same chemical and physical laws regardless of whether they are a part of the living or nonliving world.

**Approximate
Percentage
of Elements
in Living
Organisms
(Humans)
Compared
to the**

Nonliving World			
Element	Life (Humans)	Atmosphere	Earth's Crust
Oxygen (O)	65%	21%	46%
Carbon (C)	18%	trace	trace
Hydrogen (H)	10%	trace	0.1%
Nitrogen (N)	3%	78%	trace

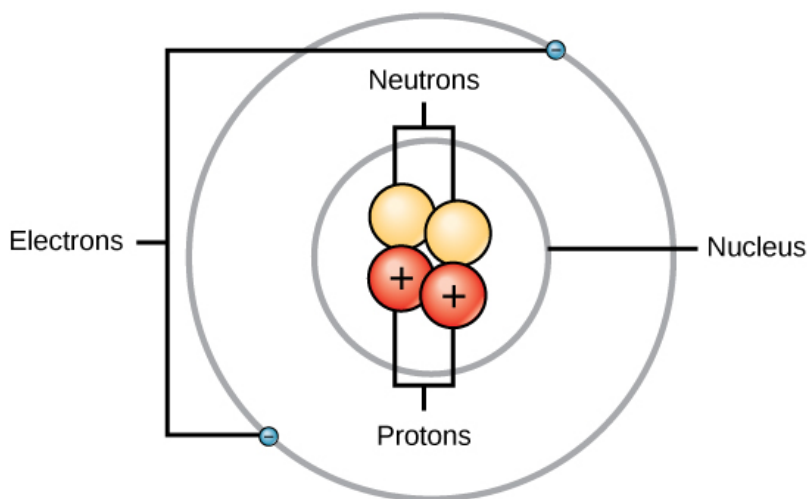
Elements, such as helium, depicted here, are made up of atoms. Atoms are made up of protons and neutrons located within the nucleus, with electrons in orbitals surrounding the nucleus.

The Structure of the Atom

To understand how elements come together, we must first discuss the element's smallest component or building block, the atom. An **atom** is the smallest unit of matter that retains all of the element's chemical properties. For example, one gold atom has all of the properties of gold in that it is a solid metal at room temperature. A gold coin is simply a very large number of gold atoms molded into the shape of a coin and contains small amounts of other elements known as impurities. We cannot break down gold atoms into anything smaller while still retaining the properties of gold.

An atom is composed of two regions: the **nucleus**,

which is in the atom's center and contains protons and neutrons. The atom's outermost region holds its electrons in orbit around the nucleus, as [\[link\]](#) illustrates. Atoms contain protons, electrons, and neutrons, among other subatomic particles. The only exception is hydrogen (H), which is made of one proton and one electron with no neutrons.



Protons and neutrons have approximately the same mass, about 1.67×10^{-24} grams. Scientists arbitrarily define this amount of mass as one atomic mass unit (amu) or one Dalton, as [\[link\]](#) shows. Although similar in mass, protons and neutrons differ in their electric charge. A **proton** is positively charged; whereas, a **neutron** is uncharged. Therefore, the number of neutrons in an atom contributes significantly to its mass, but not to its charge. **Electrons** are much smaller in mass than protons, weighing only 9.11×10^{-28} grams, or about 1/1800 of an atomic mass unit. Hence, they

do not contribute much to an element's overall atomic mass. Therefore, when considering atomic mass, it is customary to ignore the mass of any electrons and calculate the atom's mass based on the number of protons and neutrons alone. Although not significant contributors to mass, electrons do contribute greatly to the atom's charge, as each electron has a negative charge equal to the proton's positive charge. In uncharged, neutral atoms, the number of electrons orbiting the nucleus is equal to the number of protons inside the nucleus. In these atoms, the positive and negative charges cancel each other out, leading to an atom with no net charge.

Accounting for the sizes of protons, neutrons, and electrons, most of the atom's volume—greater than 99 percent—is empty space. With all this empty space, one might ask why so-called solid objects do not just pass through one another. The reason they do not is that the electrons that surround all atoms are negatively charged and negative charges repel each other.

Protons, Neutrons, and Electrons					

	Charge	Mass (amu)	Location
Proton	+1	1	nucleus
Neutron	0	1	nucleus
Electron	-1	0	orbitals

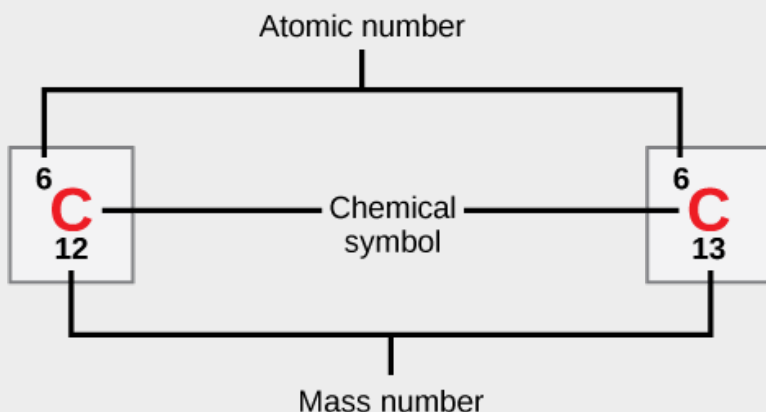
Atomic Number and Mass

Atoms of each element contain a characteristic number of protons and electrons. The number of protons determines an element's **atomic number**, which scientists use to distinguish one element from another. The number of neutrons is variable, resulting in isotopes, which are different forms of the same atom that vary only in the number of neutrons they possess. Together, the number of protons and neutrons determine an element's **mass number**, as [\[link\]](#) illustrates. Note that we disregard the small contribution of mass from electrons in calculating the mass number. We can use this approximation of mass to easily calculate how many neutrons an element has by simply subtracting the number of protons from the mass number. Since an element's isotopes will have slightly different mass numbers, scientists also determine the **atomic mass**, which is the calculated mean of the mass number for its naturally occurring isotopes. Often, the resulting number contains a fraction. For example, the atomic mass of chlorine (Cl) is 35.45 because chlorine is composed of several isotopes, some (the

majority) with atomic mass 35 (17 protons and 18 neutrons) and some with atomic mass 37 (17 protons and 20 neutrons).

Visual Connection

Carbon has an atomic number of six, and two stable isotopes with mass numbers of twelve and thirteen, respectively. Its relative atomic mass is 12.011



How many neutrons do carbon-12 and carbon-13 have, respectively?

Isotopes

Isotopes are different forms of an element that have the same number of protons but a different number

of neutrons. Some elements—such as carbon, potassium, and uranium—have naturally occurring isotopes. Carbon-12 contains six protons, six neutrons, and six electrons; therefore, it has a mass number of 12 (six protons and six neutrons).

Carbon-14 contains six protons, eight neutrons, and six electrons; its atomic mass is 14 (six protons and eight neutrons). These two alternate forms of carbon are isotopes. Some isotopes may emit neutrons, protons, and electrons, and attain a more stable atomic configuration (lower level of potential energy); these are radioactive isotopes, or **radioisotopes**. Radioactive decay (carbon-14 decaying to eventually become nitrogen-14) describes the energy loss that occurs when an unstable atom's nucleus releases radiation.

Evolution Connection

Carbon Dating

Carbon is normally present in the atmosphere in the form of gaseous compounds like carbon dioxide and methane. Carbon-14 (^{14}C) is a naturally occurring radioisotope that is created in the atmosphere from atmospheric ^{14}N (nitrogen) by the addition of a neutron and the loss of a proton because of cosmic rays. This is a continuous process, so more ^{14}C is always being created. As a living organism incorporates ^{14}C initially as carbon dioxide fixed in the process of photosynthesis, the

relative amount of ^{14}C in its body is equal to the concentration of ^{14}C in the atmosphere. When an organism dies, it is no longer ingesting ^{14}C , so the ratio between ^{14}C and ^{12}C will decline as ^{14}C decays gradually to ^{14}N by a process called beta decay—electrons or positrons emission. This decay emits energy in a slow process.

After approximately 5,730 years, half of the starting concentration of ^{14}C will convert back to ^{14}N . We call the time it takes for half of the original concentration of an isotope to decay back to its more stable form its half-life. Because the half-life of ^{14}C is long, scientists use it to date formerly living objects such as old bones or wood. Comparing the ratio of the ^{14}C concentration in an object to the amount of ^{14}C in the atmosphere, scientists can determine the amount of the isotope that has not yet decayed. On the basis of this amount, [\[link\]](#) shows that we can calculate the age of the material, such as the pygmy mammoth, with accuracy if it is not much older than about 50,000 years. Other elements have isotopes with different half lives. For example, ^{40}K (potassium-40) has a half-life of 1.25 billion years, and ^{235}U (Uranium 235) has a half-life of about 700 million years. Through the use of radiometric dating, scientists can study the age of fossils or other remains of extinct organisms to understand how organisms have evolved from earlier species.

Scientists can determine the age of carbon-containing remains less than about 50,000 years

old, such as this pygmy mammoth, using carbon dating. (credit: Bill Faulkner, NPS)



Link to Learning

To learn more about atoms, isotopes, and how to tell one isotope from another, run the simulation.

https://openstax.org/l/atoms_isotopes

The periodic table shows each element's atomic mass and atomic number. The atomic number appears above the symbol for the element and the approximate atomic mass appears below it.

The Periodic Table

The **periodic table** organizes and displays different elements. Devised by Russian chemist Dmitri Mendeleev (1834–1907) in 1869, the table groups elements that, although unique, share certain chemical properties with other elements. The properties of elements are responsible for their physical state at room temperature: they may be gases, solids, or liquids. Elements also have specific **chemical reactivity**, the ability to combine and to chemically bond with each other.

In the periodic table in [\[link\]](#), the elements are organized and displayed according to their atomic number and are arranged in a series of rows and columns based on shared chemical and physical properties. In addition to providing the atomic number for each element, the periodic table also displays the element's atomic mass. Looking at carbon, for example, its symbol (C) and name appear, as well as its atomic number of six (in the upper left-hand corner) and its atomic mass of 12.11.

Period

Group

Periodic Table of the Elements

1	1																2																
1	H																2																
1	1.008																4.003																
1	hydrogen																helium																
2	2																3																
2	Li																Be																
2	6.94																9.012																
2	lithium																beryllium																
3	3																4	5	6	7	8	9	10	11	12								
3	Na																Mg	Al	Si	P	S	Cl	Ar										
3	22.99																24.31	26.98	28.09	30.97	32.06	35.45	39.95										
3	sodium																magnesium	aluminum	silicon	phosphorus	sulfur	chlorine	argon										
4	19																20	21	22	23	24	25	26	27	28	29	30						
4	K																Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn						
4	39.10																40.08	44.96	47.87	50.94	52.00	54.94	55.85	58.93	58.69	63.55	65.38						
4	potassium																calcium	scandium	titanium	vanadium	chromium	manganese	iron	cobalt	nickel	copper	zinc						
5	37																38	39	40	41	42	43	44	45	46	47	48	49	50				
5	Rb																Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn				
5	85.47																87.62	88.91	91.22	92.91	95.95	97	101.1	102.9	106.4	107.9	112.4	114.8	118.7				
5	rubidium																strontium	yttrium	zirconium	niobium	molybdenum	technetium	ruthenium	rhodium	palladium	silver	cadmium	indium	tin				
6	55																56	57-71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
6	Cs																Ba	La-Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
6	132.9																137.3	lanthanum series	178.5	180.9	183.8	186.2	190.2	192.2	195.1	197.0	200.6	204.4	207.2	209.0	[209]	[210]	[222]
6	cesium																barium	lanthanum series	hafnium	tantalum	tungsten	rhenium	osmium	iridium	platinum	gold	mercury	thallium	lead	bismuth	polonium	astatine	radon
7	87																88	89-103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118
7	Fr																Ra	Ac-Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Nh	Fl	Mc	Lv	Ts	Og
7	[223]																[226]	actinium series	[261]	[270]	[271]	[270]	[277]	[276]	[281]	[284]	[285]	[285]	[289]	[288]	[293]	[294]	[294]
7	francium																radium	actinium series	rutherfordium	dubnium	seaborgium	bohrium	hassium	meitnerium	darmstadtium	roentgenium	copernicium	nihonium	flerovium	moscovium	livermorium	tennessine	oganesson
	57																58	59	60	61	62	63	64	65	66	67	68	69	70	71			
	La																Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu			
	138.9																140.1	140.9	144.2	[145]	150.4	152.0	157.3	158.9	162.5	164.9	167.3	168.9	173.1	175.0			
	lanthanum																cerium	praseodymium	neodymium	promethium	samarium	europium	gadolinium	terbium	dysprosium	holmium	erbium	thulium	ytterbium	lutetium			
	89																90	91	92	93	94	95	96	97	98	99	100	101	102	103			
	Ac																Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr			
	[227]																232.0	231.0	238.0	[237]	[244]	[243]	[247]	[247]	[251]	[252]	[257]	[258]	[259]	[262]			
	actinium																thorium	protactinium	uranium	neptunium	plutonium	americium	curium	berkelium	californium	einsteinium	fermium	mendelevium	nobelium	lawrencium			

Atomic number → 1

Symbol → **H**

Atomic mass → 1.008

Name → hydrogen

Color Code

Metal

Metalloid

Nonmetal

Solid

Liquid

Gas

The periodic table groups elements according to chemical properties. Scientists base the differences in chemical reactivity between the elements on the number and spatial distribution of an atom's electrons. Atoms that chemically react and bond to each other form molecules. **Molecules** are simply two or more atoms chemically bonded together. Logically, when two atoms chemically bond to form a molecule, their electrons, which form the outermost of each atom, come together first as the atoms form a chemical bond.

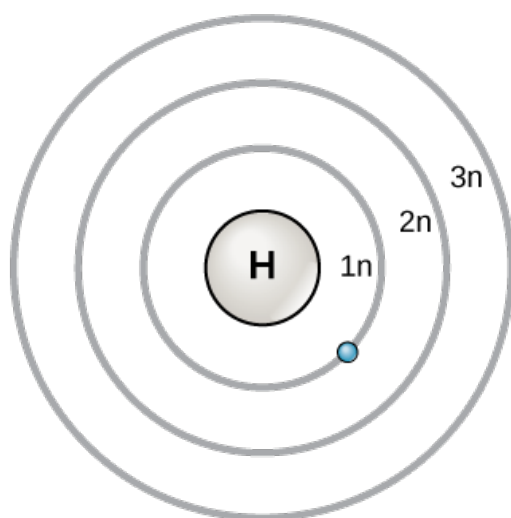
In 1913, Niels Bohrs developed the Bohr model in which electrons exist within principal shells. An electron normally exists in the lowest energy shell available, which is the one closest to the nucleus. Energy from a photon of light can bump it up to a

higher energy shell, but this situation is unstable, and the electron quickly decays back to the ground state. In the process, it releases a photon of light.

Electron Shells and the Bohr Model

Note that there is a connection between the number of protons in an element, the atomic number that distinguishes one element from another, and the number of electrons it has. In all electrically neutral atoms, the number of electrons is the same as the number of protons. Thus, each element, at least when electrically neutral, has a characteristic number of electrons equal to its atomic number.

In 1913, Danish scientist Niels Bohr (1885–1962) developed an early model of the atom. The Bohr model shows the atom as a central nucleus containing protons and neutrons, with the electrons in circular **orbitals** at specific distances from the nucleus, as [\[link\]](#) illustrates. These orbits form electron shells or energy levels, which are a way of visualizing the number of electrons in the outermost shells. These energy levels are designated by a number and the symbol “n.” For example, 1n represents the first energy level located closest to the nucleus.





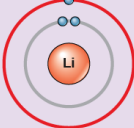
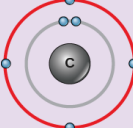
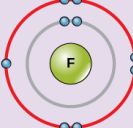
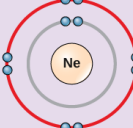

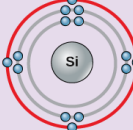
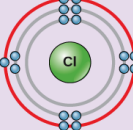
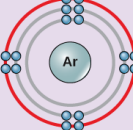
Electrons fill orbitals in a consistent order: they first fill the orbitals closest to the nucleus, then they continue to fill orbitals of increasing energy further from the nucleus. If there are multiple orbitals of equal energy, they fill with one electron in each energy level before adding a second electron. The electrons of the outermost energy level determine the atom's energetic stability and its tendency to form chemical bonds with other atoms to form molecules.

Under standard conditions, atoms fill the inner shells first, often resulting in a variable number of electrons in the outermost shell. The innermost shell has a maximum of two electrons but the next two

electron shells can each have a maximum of eight electrons. This is known as the **octet rule**, which states, with the exception of the innermost shell, that atoms are more stable energetically when they have eight electrons in their **valence shell**, the outermost electron shell. [\[link\]](#) shows examples of some neutral atoms and their electron configurations. Notice that in [\[link\]](#), helium has a complete outer electron shell, with two electrons filling its first and only shell. Similarly, neon has a complete outer $2n$ shell containing eight electrons. In contrast, chlorine and sodium have seven and one in their outer shells, respectively, but theoretically they would be more energetically stable if they followed the octet rule and had eight.

Visual Connection

Bohr diagrams indicate how many electrons fill each principal shell. Group 18 elements (helium, neon, and argon) have a full outer, or valence, shell. A full valence shell is the most stable electron configuration. Elements in other groups have partially filled valence shells and gain or lose electrons to achieve a stable electron configuration.

	Group 1	Group 14	Group 17	Group 18
Period 1 (1n is filling)				
Period 2 (2n is filling)				
Period 3 (3n is filling)				

An atom may give, take, or share electrons with another atom to achieve a full valence shell, the most stable electron configuration. Looking at this figure, how many electrons do elements in group 1 need to lose in order to achieve a stable electron configuration? How many electrons do elements in groups 14 and 17 need to gain to achieve a stable configuration?

Understanding that the periodic table's organization is based on the total number of protons (and electrons) helps us know how electrons distribute themselves among the shells. The periodic table is arranged in columns and rows based on the number of electrons and their location. Examine more closely some of the elements in the table's far right column in [\[link\]](#). The group 18 atoms helium (He), neon (Ne), and argon (Ar) all have filled outer

electron shells, making it unnecessary for them to share electrons with other atoms to attain stability. They are highly stable as single atoms. Because they are non reactive, scientists coin them **inert** (or **noble gases**). Compare this to the group 1 elements in the left-hand column. These elements, including hydrogen (H), lithium (Li), and sodium (Na), all have one electron in their outermost shells. That means that they can achieve a stable configuration and a filled outer shell by donating or sharing one electron with another atom or a molecule such as water. Hydrogen will donate or share its electron to achieve this configuration, while lithium and sodium will donate their electron to become stable. As a result of losing a negatively charged electron, they become positively charged **ions**. Group 17 elements, including fluorine and chlorine, have seven electrons in their outmost shells, so they tend to fill this shell with an electron from other atoms or molecules, making them negatively charged ions. Group 14 elements, of which carbon is the most important to living systems, have four electrons in their outer shell allowing them to make several covalent bonds (discussed below) with other atoms. Thus, the periodic table's columns represent the potential shared state of these elements' outer electron shells that is responsible for their similar chemical characteristics.

The s subshells are shaped like spheres. Both the $1n$ and $2n$ principal shells have an s orbital, but the size of the sphere is larger in the $2n$ orbital. Each sphere

is a single orbital. Three dumbbell-shaped orbitals comprise p subshells. Principal shell $2n$ has a p subshell, but shell 1 does not.

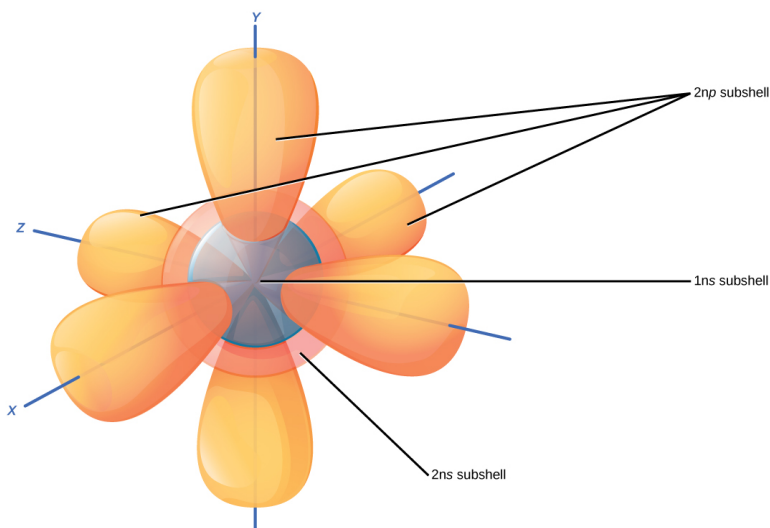
Electron Orbitals

Although useful to explain the reactivity and chemical bonding of certain elements, the Bohr model does not accurately reflect how electrons spatially distribute themselves around the nucleus. They do not circle the nucleus like the earth orbits the sun, but we find them in **electron orbitals**.

These relatively complex shapes result from the fact that electrons behave not just like particles, but also like waves. Mathematical equations from quantum mechanics, which scientists call wave functions, can predict within a certain level of probability where an electron might be at any given time. Scientists call the area where an electron is most likely to be found its orbital.

Recall that the Bohr model depicts an atom's electron shell configuration. Within each electron shell are subshells, and each subshell has a specified number of orbitals containing electrons. While it is impossible to calculate exactly an electron's location, scientists know that it is most probably located within its orbital path. The letter s , p , d , and f designate the subshells. The s subshell is spherical in shape and has one orbital. Principal shell $1n$ has only a single s orbital, which can hold two electrons.

Principal shell $2n$ has one s and one p subshell, and can hold a total of eight electrons. The p subshell has three dumbbell-shaped orbitals, as [\[link\]](#) illustrates. Subshells d and f have more complex shapes and contain five and seven orbitals, respectively. We do not show these in the illustration. Principal shell $3n$ has s , p , and d subshells and can hold 18 electrons. Principal shell $4n$ has s , p , d and f orbitals and can hold 32 electrons. Moving away from the nucleus, the number of electrons and orbitals in the energy levels increases. Progressing from one atom to the next in the periodic table, we can determine the electron structure by fitting an extra electron into the next available orbital.



The closest orbital to the nucleus, the $1s$ orbital, can hold up to two electrons. This orbital is equivalent to the Bohr model's innermost electron shell. Scientists call it the $1s$ orbital because it is spherical

around the nucleus. The $1s$ orbital is the closest orbital to the nucleus, and it is always filled first, before any other orbital fills. Hydrogen has one electron; therefore, it occupies only one spot within the $1s$ orbital. We designate this as $1s_1$, where the superscripted 1 refers to the one electron within the $1s$ orbital. Helium has two electrons; therefore, it can completely fill the $1s$ orbital with its two electrons. We designate this as $1s_2$, referring to the two electrons of helium in the $1s$ orbital. On the periodic table [\[link\]](#), hydrogen and helium are the only two elements in the first row (period). This is because they only have electrons in their first shell, the $1s$ orbital. Hydrogen and helium are the only two elements that have the $1s$ and no other electron orbitals in the electrically neutral state.

The second electron shell may contain eight electrons. This shell contains another spherical s orbital and three “dumbbell” shaped p orbitals, each of which can hold two electrons, as [\[link\]](#) shows. After the $1s$ orbital fills, the second electron shell fills, first filling its $2s$ orbital and then its three p orbitals. When filling the p orbitals, each takes a single electron. Once each p orbital has an electron, it may add a second. Lithium (Li) contains three electrons that occupy the first and second shells. Two electrons fill the $1s$ orbital, and the third electron then fills the $2s$ orbital. Its **electron configuration** is $1s_2 2s_1$. Neon (Ne), alternatively, has a total of ten electrons: two are in its innermost

1s orbital and eight fill its second shell (two each in the 2s and three *p* orbitals). Thus it is an inert gas and energetically stable as a single atom that will rarely form a chemical bond with other atoms. Larger elements have additional orbitals, comprising the third electron shell. While the concepts of electron shells and orbitals are closely related, orbitals provide a more accurate depiction of an atom's electron configuration because the orbital model specifies the different shapes and special orientations of all the places that electrons may occupy.

Link to Learning

Watch this visual animation to see the spatial arrangement of the *p* and *s* orbitals.

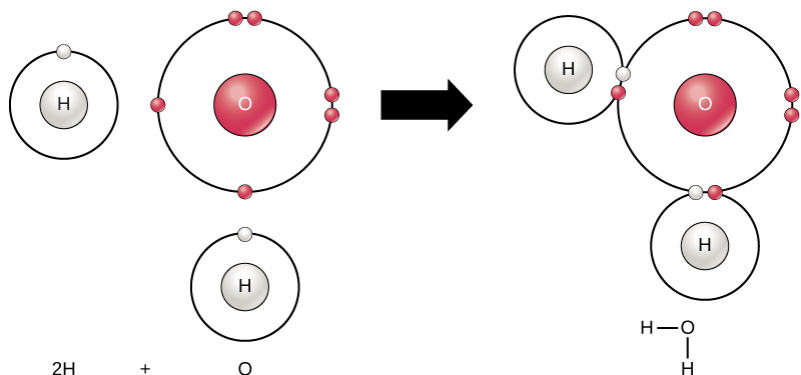
<https://www.openstax.org/l/orbitals>

Two or more atoms may bond with each other to form a molecule. When two hydrogens and an oxygen share electrons via covalent bonds it forms a water molecule. A double bond joins the oxygen atoms in an O₂ molecule.

Chemical Reactions and Molecules

All elements are most stable when their outermost

shell is filled with electrons according to the octet rule. This is because it is energetically favorable for atoms to be in that configuration and it makes them stable. However, since not all elements have enough electrons to fill their outermost shells, atoms form **chemical bonds** with other atoms thereby obtaining the electrons they need to attain a stable electron configuration. When two or more atoms chemically bond with each other, the resultant chemical structure is a molecule. The familiar water molecule, H_2O , consists of two hydrogen atoms and one oxygen atom. These bond together to form water, as [\[link\]](#) illustrates. Atoms can form molecules by donating, accepting, or sharing electrons to fill their outer shells.



Chemical reactions occur when two or more atoms bond together to form molecules or when bonded atoms break apart. Scientists call the substances used in the beginning of a chemical reaction **reactants** (usually on the left side of a chemical equation), and we call the substances at the end of the reaction **products** (usually on the right side of a

chemical equation). We typically draw an arrow between the reactants and products to indicate the chemical reaction's direction. This direction is not always a “one-way street.” To create the water molecule above, the chemical equation would be:

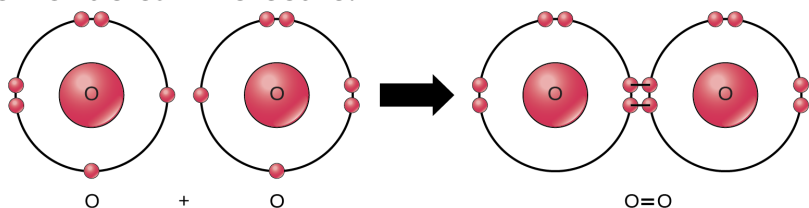
$$2\text{H} + \text{O} \rightarrow \text{H}_2\text{O}$$

An example of a simple chemical reaction is breaking down hydrogen peroxide molecules, each of which consists of two hydrogen atoms bonded to two oxygen atoms (H_2O_2). The reactant hydrogen peroxide breaks down into water, containing one oxygen atom bound to two hydrogen atoms (H_2O), and oxygen, which consists of two bonded oxygen atoms (O_2). In the equation below, the reaction includes two hydrogen peroxide molecules and two water molecules. This is an example of a **balanced chemical equation**, wherein each element's number of atoms is the same on each side of the equation. According to the law of conservation of matter, the number of atoms before and after a chemical reaction should be equal, such that no atoms are, under normal circumstances, created or destroyed.

$$2\text{H}_2\text{O}_2 \text{ (hydrogen peroxide)} \rightarrow 2\text{H}_2\text{O} \text{ (water)} + \text{O}_2 \text{ (oxygen)}$$

Even though all of the reactants and products of this reaction are molecules (each atom remains bonded to at least one other atom), in this reaction only hydrogen peroxide and water are representatives of **compounds**: they contain atoms of more than one

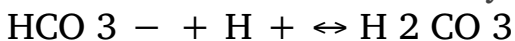
type of element. Molecular oxygen, alternatively, as [\[link\]](#) shows, consists of two doubly bonded oxygen atoms and is not classified as a compound but as a homonuclear molecule.



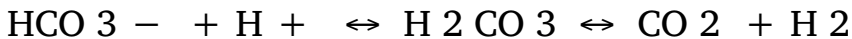
Some chemical reactions, such as the one above, can proceed in one direction until they expend all the reactants. The equations that describe these reactions contain a unidirectional arrow and are **irreversible**. **Reversible reactions** are those that can go in either direction. In reversible reactions, reactants turn into products, but when the product's concentration goes beyond a certain threshold (characteristic of the particular reaction), some of these products convert back into reactants. At this point, product and reactant designations reverse. This back and forth continues until a certain relative balance between reactants and products occurs—a state called **equilibrium**. A chemical equation with a double headed arrow pointing towards both the reactants and products often denote these reversible reaction situations.

For example, in human blood, excess hydrogen ions (H⁺) bind to bicarbonate ions (HCO₃⁻) forming an equilibrium state with carbonic acid (H₂CO₃). If we added carbonic acid to this system, some of it would

convert to bicarbonate and hydrogen ions.



However, biological reactions rarely obtain equilibrium because the concentrations of the reactants or products or both are constantly changing, often with one reaction's product a reactant for another. To return to the example of excess hydrogen ions in the blood, forming carbonic acid will be the reaction's major direction. However, the carbonic acid can also leave the body as carbon dioxide gas (via exhalation) instead of converting back to bicarbonate ion, thus driving the reaction to the right by the **law of mass action**. These reactions are important for maintaining homeostasis in our blood.



In the formation of an ionic compound, metals lose electrons and nonmetals gain electrons to achieve an octet.

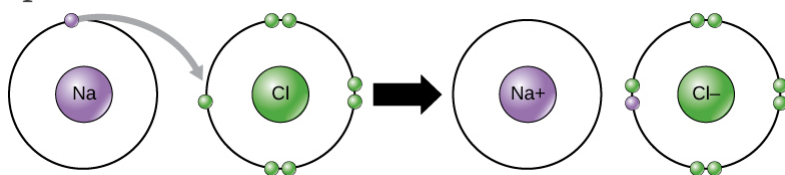
Ions and Ionic Bonds

Some atoms are more stable when they gain or lose an electron (or possibly two) and form ions. This fills their outermost electron shell and makes them energetically more stable. Because the number of electrons does not equal the number of protons, each ion has a net charge. **Cations** are positive ions that form by losing electrons. Negative ions form by gaining electrons, which we call anions. We

designate **anions** by their elemental name and change the ending to “-ide”, thus the anion of chlorine is chloride, and the anion of sulfur is sulfide.

Scientists refer to this movement of electrons from one element to another as **electron transfer**. As [\[link\]](#) illustrates, sodium (Na) only has one electron in its outer electron shell. It takes less energy for sodium to donate that one electron than it does to accept seven more electrons to fill the outer shell. If sodium loses an electron, it now has 11 protons, 11 neutrons, and only 10 electrons, leaving it with an overall charge of $+1$. We now refer to it as a sodium ion. Chlorine (Cl) in its lowest energy state (called the ground state) has seven electrons in its outer shell. Again, it is more energy-efficient for chlorine to gain one electron than to lose seven. Therefore, it tends to gain an electron to create an ion with 17 protons, 17 neutrons, and 18 electrons, giving it a net negative (-1) charge. We now refer to it as a chloride ion. In this example, sodium will donate its one electron to empty its shell, and chlorine will accept that electron to fill its shell. Both ions now satisfy the octet rule and have complete outermost shells. Because the number of electrons is no longer equal to the number of protons, each is now an ion and has a $+1$ (sodium cation) or -1 (chloride anion) charge. Note that these transactions can normally only take place simultaneously: in order for a sodium atom to lose

an electron, it must be in the presence of a suitable recipient like a chlorine atom.



Ionic bonds form between ions with opposite charges. For instance, positively charged sodium ions and negatively charged chloride ions bond together to make crystals of sodium chloride, or table salt, creating a crystalline molecule with zero net charge.

Physiologists refer to certain salts as **electrolytes** (including sodium, potassium, and calcium), ions necessary for nerve impulse conduction, muscle contractions, and water balance. Many sports drinks and dietary supplements provide these ions to replace those lost from the body via sweating during exercise.

Whether a molecule is polar or nonpolar depends both on bond type and molecular shape. Both water and carbon dioxide have polar covalent bonds, but carbon dioxide is linear, so the partial charges on the molecule cancel each other out.

Covalent Bonds and Other Bonds and Interactions

Another way to satisfy the octet rule is by sharing

electrons between atoms to form **covalent bonds**. These bonds are stronger and much more common than ionic bonds in the molecules of living organisms. We commonly find covalent bonds in carbon-based organic molecules, such as our DNA and proteins. We also find covalent bonds in inorganic molecules like H_2O , CO_2 , and O_2 . The bonds may share one, two, or three pairs of electrons, making single, double, and triple bonds, respectively. The more covalent bonds between two atoms, the stronger their connection. Thus, triple bonds are the strongest.

The strength of different levels of covalent bonding is one of the main reasons living organisms have a difficult time in acquiring nitrogen for use in constructing their molecules, even though molecular nitrogen, N_2 , is the most abundant gas in the atmosphere. Molecular nitrogen consists of two nitrogen atoms triple bonded to each other and, as with all molecules, sharing these three pairs of electrons between the two nitrogen atoms allows for filling their outer electron shells, making the molecule more stable than the individual nitrogen atoms. This strong triple bond makes it difficult for living systems to break apart this nitrogen in order to use it as constituents of proteins and DNA.

Forming water molecules provides an example of covalent bonding. Covalent bonds bind the hydrogen and oxygen atoms that combine to form

water molecules as [\[link\]](#) shows. The electron from the hydrogen splits its time between the hydrogen atoms' incomplete outer shell and the oxygen atoms' incomplete outer shell. To completely fill the oxygen's outer shell, which has six electrons but which would be more stable with eight, two electrons (one from each hydrogen atom) are needed: hence, the well-known formula H_2O . The two elements share the electrons to fill the outer shell of each, making both elements more stable.

Link to Learning

View this short video to see an animation of ionic and covalent bonding.

https://www.openstax.org/l/ionic_covalent

Polar Covalent Bonds

There are two types of covalent bonds: polar and nonpolar. In a **polar covalent bond**, [\[link\]](#) shows atoms unequally share the electrons and are attracted more to one nucleus than the other. Because of the unequal electron distribution between the atoms of different elements, a slightly positive ($\delta+$) or slightly negative ($\delta-$) charge develops. This partial charge is an important property of water and accounts for many of its

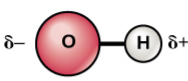
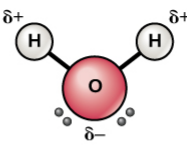
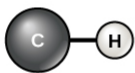
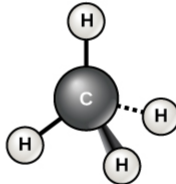
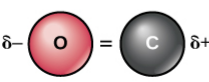
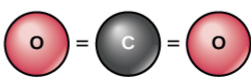
characteristics.

Water is a polar molecule, with the hydrogen atoms acquiring a partial positive charge and the oxygen a partial negative charge. This occurs because the oxygen atom's nucleus is more attractive to the hydrogen atoms' electrons than the hydrogen nucleus is to the oxygen's electrons. Thus, oxygen has a higher **electronegativity** than hydrogen and the shared electrons spend more time near the oxygen nucleus than the hydrogen atoms' nucleus, giving the oxygen and hydrogen atoms slightly negative and positive charges, respectively. Another way of stating this is that the probability of finding a shared electron near an oxygen nucleus is more likely than finding it near a hydrogen nucleus. Either way, the atom's relative electronegativity contributes to developing partial charges whenever one element is significantly more electronegative than the other, and the charges that these polar bonds generate may then be used to form hydrogen bonds based on the attraction of opposite partial charges. (Hydrogen bonds, which we discuss in detail below, are weak bonds between slightly positively charged hydrogen atoms to slightly negatively charged atoms in other molecules.) Since macromolecules often have atoms within them that differ in electronegativity, polar bonds are often present in organic molecules.

Nonpolar Covalent Bonds

Nonpolar covalent bonds form between two atoms of the same element or between different elements that share electrons equally. For example, molecular oxygen (O₂) is nonpolar because the electrons distribute equally between the two oxygen atoms.

[\[link\]](#) also shows another example of a nonpolar covalent bond—methane (CH₄). Carbon has four electrons in its outermost shell and needs four more to fill it. It obtains these four from four hydrogen atoms, each atom providing one, making a stable outer shell of eight electrons. Carbon and hydrogen do not have the same electronegativity but are similar; thus, nonpolar bonds form. The hydrogen atoms each need one electron for their outermost shell, which is filled when it contains two electrons. These elements share the electrons equally among the carbons and the hydrogen atoms, creating a nonpolar covalent molecule.

	Bond type	Molecular shape	Molecular type
Water	 Polar covalent	 Bent	Polar
Methane	 Nonpolar covalent	 Tetrahedral	Nonpolar
Carbon dioxide	 Polar covalent	 Linear	Nonpolar

Hydrogen Bonds and Van Der Waals Interactions

Ionic and covalent bonds between elements require energy to break. Ionic bonds are not as strong as covalent, which determines their behavior in biological systems. However, not all bonds are ionic or covalent bonds. Weaker bonds can also form between molecules. Two weak bonds that occur frequently are hydrogen bonds and van der Waals interactions. Without these two types of bonds, life as we know it would not exist. Hydrogen bonds provide many of the critical, life-sustaining properties of water and also stabilize the structures of proteins and DNA, the building block of cells.

When polar covalent bonds containing hydrogen form, the hydrogen in that bond has a slightly positive charge because hydrogen's electron is pulled more strongly toward the other element and away from the hydrogen. Because the hydrogen is slightly positive, it will be attracted to neighboring negative charges. When this happens, a weak interaction occurs between the hydrogen's $\delta+$ from one molecule and another molecule's $\delta-$ charge on the more electronegative atoms, usually oxygen or nitrogen, or within the same molecule. Scientists call this interaction a **hydrogen bond**. This type of bond is common and occurs regularly between water molecules. Individual hydrogen bonds are weak and easily broken; however, they occur in very large numbers in water and in organic polymers, creating a major force in combination. Hydrogen bonds are also responsible for zipping together the DNA double helix.

Like hydrogen bonds, **van der Waals interactions** are weak attractions or interactions between molecules. Van der Waals attractions can occur between any two or more molecules and are dependent on slight fluctuations of the electron densities, which are not always symmetrical around an atom. For these attractions to happen, the molecules need to be very close to one another. These bonds—along with ionic, covalent, and hydrogen bonds—contribute to the proteins' three-dimensional structure in our cells that is necessary

for their proper function.

Career Connection

Pharmaceutical Chemist

Pharmaceutical chemists are responsible for developing new drugs and trying to determine the mode of action of both old and new drugs. They are involved in every step of the drug development process. We can find drugs in the natural environment or we can synthesize them in the laboratory. In many cases, chemists chemically change potential drugs from nature chemically in the laboratory to make them safer and more effective, and sometimes synthetic versions of drugs substitute for the version we find in nature. After a drug's initial discovery or synthesis, the chemist then develops the drug, perhaps chemically altering it, testing it to see if it is toxic, and then designing methods for efficient large-scale production. Then, the process of approving the drug for human use begins. In the United States, the Food and Drug Administration (FDA) handles drug approval. This involves a series of large-scale experiments using human subjects to ensure the drug is not harmful and effectively treats the condition for which it is intended. This process often takes several years and requires the participation of physicians and scientists, in addition to chemists, to complete testing and gain

approval.

An example of a drug that was originally discovered in a living organism is Paclitaxel (Taxol), an anti-cancer drug used to treat breast cancer. This drug was discovered in the bark of the pacific yew tree. Another example is aspirin, originally isolated from willow tree bark. Finding drugs often means testing hundreds of samples of plants, fungi, and other forms of life to see if they contain any biologically active compounds.

Sometimes, traditional medicine can give modern medicine clues as to where to find an active compound. For example, mankind has used willow bark to make medicine for thousands of years, dating back to ancient Egypt. However, it was not until the late 1800s that scientists and pharmaceutical companies purified and marketed the aspirin molecule, acetylsalicylic acid, for human use.

Occasionally, drugs developed for one use have unforeseen effects that allow usage in other, unrelated ways. For example, scientists originally developed the drug minoxidil (Rogaine) to treat high blood pressure. When tested on humans, researchers noticed that individuals taking the drug would grow new hair. Eventually the pharmaceutical company marketed the drug to men and women with baldness to restore lost hair. A pharmaceutical chemist's career may involve detective work, experimentation, and drug development, all with the goal of making human

beings healthier.

Section Summary

Matter is anything that occupies space and has mass. It is comprised of elements. All of the 98 elements that occur naturally have unique qualities that allow them to combine in various ways to create molecules, which in turn combine to form cells, tissues, organ systems, and organisms. Atoms, which consist of protons, neutrons, and electrons, are the smallest units of an element that retain all of the properties of that element. Electrons can transfer, share, or cause charge disparities between atoms to create bonds, including ionic, covalent, and hydrogen bonds, as well as van der Waals interactions.

Visual Connection Questions

[\[link\]](#) How many neutrons do carbon-12 and carbon-13 have, respectively?

[\[link\]](#) Carbon-12 has six neutrons. Carbon-13 has seven neutrons.

[\[link\]](#) An atom may give, take, or share electrons with another atom to achieve a full valence shell, the most stable electron configuration. Looking at this figure, how many electrons do elements in group 1 need to lose in order to achieve a stable electron configuration? How many electrons do elements in groups 14 and 17 need to gain to achieve a stable configuration?

[\[link\]](#) Elements in group 1 need to lose one electron to achieve a stable electron configuration. Elements in groups 14 and 17 need to gain four and one electrons, respectively, to achieve a stable configuration.

Review Questions

If xenon has an atomic number of 54 and a mass number of 108, how many neutrons does it have?

1. 54

2. 27
 3. 100
 4. 108
-

A

Atoms that vary in the number of neutrons found in their nuclei are called ____.

1. ions
 2. neutrons
 3. neutral atoms
 4. isotopes
-

D

Potassium has an atomic number of 19. What is its electron configuration?

1. shells 1 and 2 are full, and shell 3 has nine electrons
 2. shells 1, 2 and 3 are full and shell 4 has three electrons
 3. shells 1, 2 and 3 are full and shell 4 has one electron
 4. shells 1, 2 and 3 are full and no other electrons are present
-

C

Which type of bond represents a weak chemical bond?

1. hydrogen bond
2. atomic bond
3. covalent bond
4. nonpolar covalent bond

A

Critical Thinking Questions

What makes ionic bonds different from covalent bonds?

Ionic bonds are created between ions. The electrons are not shared between the atoms, but rather are associated more with one ion than the other. Ionic bonds are strong bonds, but are weaker than covalent bonds, meaning it takes less energy to break an ionic bond compared with a covalent one.

Why are hydrogen bonds and van der Waals interactions necessary for cells?

Hydrogen bonds and van der Waals interactions form weak associations between different molecules or within different regions of the same molecule. They provide the structure and shape necessary for proteins and DNA within cells so that they function properly.

Glossary

anion

negative ion that is formed by an atom gaining one or more electrons

atom

the smallest unit of matter that retains all of the chemical properties of an element

atomic mass

calculated mean of the mass number for an element's isotopes

atomic number

total number of protons in an atom

balanced chemical equation

statement of a chemical reaction with the number of each type of atom equalized for

both the products and reactants

cation

positive ion that is formed by an atom losing one or more electrons

chemical bond

interaction between two or more of the same or different atoms that results in forming molecules

chemical reaction

process leading to rearranging atoms in molecules

chemical reactivity

the ability to combine and to chemically bond with each other

compound

substance composed of molecules consisting of atoms of at least two different elements

covalent bond

type of strong bond formed between two atoms of the same or different elements; forms when electrons are shared between atoms

electrolyte

ion necessary for nerve impulse conduction, muscle contractions, and water balance

electron

negatively charged subatomic particle that resides outside of the nucleus in the electron orbital; lacks functional mass and has a negative charge of -1 unit

electron configuration

arrangement of electrons in an atom's electron shell (for example, $1s^2 2s^2 2p^6$)

electron orbital

how electrons are spatially distributed surrounding the nucleus; the area where we are most likely to find an electron

electron transfer

movement of electrons from one element to another; important in creating ionic bonds

electronegativity

ability of some elements to attract electrons (often of hydrogen atoms), acquiring partial negative charges in molecules and creating partial positive charges on the hydrogen atoms

element

one of 118 unique substances that cannot break down into smaller substances; each element has unique properties and a specified number of protons

equilibrium

steady state of relative reactant and product concentration in reversible chemical reactions in a closed system

hydrogen bond

weak bond between slightly positively charged hydrogen atoms and slightly negatively charged atoms in other molecules

inert gas

(also, noble gas) element with filled outer electron shell that is unreactive with other atoms

ion

atom or chemical group that does not contain equal numbers of protons and electrons

ionic bond

chemical bond that forms between ions with opposite charges (cations and anions)

irreversible chemical reaction

chemical reaction where reactants proceed unidirectionally to form products

isotope

one or more forms of an element that have different numbers of neutrons

law of mass action

chemical law stating that the rate of a reaction is proportional to the concentration of the reacting substances

mass number

total number of protons and neutrons in an atom

matter

anything that has mass and occupies space

molecule

two or more atoms chemically bonded together

neutron

uncharged particle that resides in an atom's nucleus; has a mass of one amu

noble gas

see inert gas

nonpolar covalent bond

type of covalent bond that forms between atoms when electrons are shared equally between them

nucleus

core of an atom; contains protons and neutrons

octet rule

rule that atoms are most stable when they

hold eight electrons in their outermost shells

orbital

region surrounding the nucleus; contains electrons

periodic table

organizational chart of elements indicating each element's atomic number and atomic mass; provides key information about the elements' properties

polar covalent bond

type of covalent bond that forms as a result of unequal electron sharing, resulting in creating slightly positive and negative charged molecule regions

product

molecule that is result of chemical reaction

proton

positively charged particle that resides in the atom's nucleus; has a mass of one amu and a charge of +1

radioisotope

isotope that emits radiation comprised of subatomic particles to form more stable elements

reactant

molecule that takes part in a chemical reaction

reversible chemical reaction

chemical reaction that functions bidirectionally, where products may turn into reactants if their concentration is great enough

valence shell

outermost shell of an atom

van der Waals interaction

very weak interaction between molecules due to temporary charges attracting atoms that are very close together

Water

By the end of this section, you will be able to do the following:

- Describe the properties of water that are critical to maintaining life
- Explain why water is an excellent solvent
- Provide examples of water's cohesive and adhesive properties
- Discuss the role of acids, bases, and buffers in homeostasis

Why do scientists spend time looking for water on other planets? Why is water so important? It is because water is essential to life as we know it. Water is one of the more abundant molecules and the one most critical to life on Earth. Water comprises approximately 60–70 percent of the human body. Without it, life as we know it simply would not exist.

The polarity of the water molecule and its resulting hydrogen bonding make water a unique substance with special properties that are intimately tied to the processes of life. Life originally evolved in a watery environment, and most of an organism's cellular chemistry and metabolism occur inside the watery contents of the cell's cytoplasm. Special properties of water are its high heat capacity and heat of vaporization, its ability to dissolve polar molecules, its cohesive and adhesive properties, and

its dissociation into ions that leads to generating pH. Understanding these characteristics of water helps to elucidate its importance in maintaining life.

Oil and water do not mix. As this macro image of oil and water shows, oil does not dissolve in water but forms droplets instead. This is because it is a nonpolar compound. (credit: Gautam Dogra).

Water's Polarity

One of water's important properties is that it is composed of polar molecules: the hydrogen and oxygen within water molecules (H_2O) form polar covalent bonds. While there is no net charge to a water molecule, water's polarity creates a slightly positive charge on hydrogen and a slightly negative charge on oxygen, contributing to water's properties of attraction. Water generates charges because oxygen is more electronegative than hydrogen, making it more likely that a shared electron would be near the oxygen nucleus than the hydrogen nucleus, thus generating the partial negative charge near the **oxygen**.

As a result of water's polarity, each water molecule attracts other water molecules because of the opposite charges between water molecules, forming hydrogen bonds. Water also attracts or is attracted to other polar molecules and ions. We call a polar substance that interacts readily with or dissolves in water **hydrophilic** (hydro- = "water"; -philic =

“loving”). In contrast, nonpolar molecules such as oils and fats do not interact well with water, as [\[link\]](#) shows. A good example of this is vinegar and oil salad dressing (an acidic water solution). We call such nonpolar compounds **hydrophobic** (hydro- = “water”; -phobic = “fearing”).



Hydrogen bonding makes ice less dense than liquid water. The (a) lattice structure of ice makes it less dense than the liquid water's freely flowing molecules, enabling it to (b) float on water. (credit a: modification of work by Jane Whitney, image created using Visual Molecular Dynamics (VMD) software[\[footnote\]](#); credit b: modification of work by Carlos Ponte)W. Humphrey W., A. Dalke, and K. Schulten, “VMD—Visual Molecular Dynamics,” *Journal of Molecular Graphics* 14 (1996): 33-38.

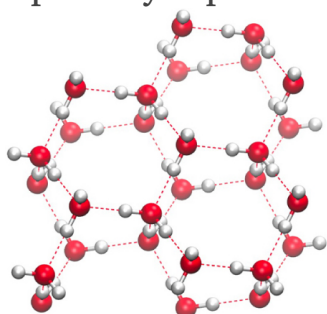
Water's States: Gas, Liquid, and Solid

The formation of hydrogen bonds is an important quality of the liquid water that is crucial to life as we know it. As water molecules make hydrogen bonds with each other, water takes on some unique chemical characteristics compared to other liquids and, since living things have a high water content, understanding these chemical features is key to understanding life. In liquid water, hydrogen bonds constantly form and break as the water molecules slide past each other. The water molecules' motion (kinetic energy) causes the bonds to break due to the heat contained in the system. When the heat rises as water boils, the water molecules' higher kinetic energy causes the hydrogen bonds to break completely and allows water molecules to escape into the air as gas (steam or water vapor). Alternatively, when water temperature reduces and water freezes, the water molecules form a crystalline structure maintained by hydrogen bonding (there is not enough energy to break the hydrogen bonds) that makes ice less dense than liquid water, a phenomenon that we do not see when other liquids solidify.

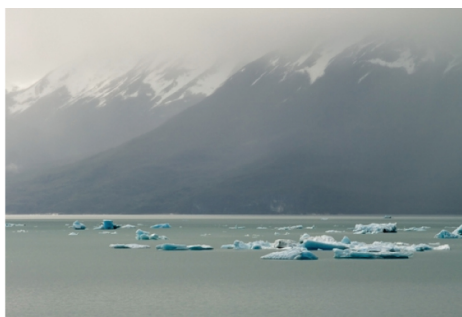
Water's lower density in its solid form is due to the way hydrogen bonds orient as they freeze: the water molecules push farther apart compared to liquid water. With most other liquids, solidification when the temperature drops includes lowering kinetic

energy between molecules, allowing them to pack even more tightly than in liquid form and giving the solid a greater density than the liquid.

The lower density of ice, as [\[link\]](#) depicts, an anomaly causes it to float at the surface of liquid water, such as in an iceberg or ice cubes in a glass of water. In lakes and ponds, ice will form on the water's surface creating an insulating barrier that protects the animals and plant life in the pond from freezing. Without this insulating ice layer, plants and animals living in the pond would freeze in the solid block of ice and could not survive. The expansion of ice relative to liquid water causes the detrimental effect of freezing on living organisms. The ice crystals that form upon freezing rupture the delicate membranes essential for living cells to function, irreversibly damaging them. Cells can only survive freezing if another liquid like glycerol temporarily replaces the water in them.



(a)



(b)

Click [here](#) to see a 3-D animation of an ice lattice structure.

Water's High Heat Capacity

Water's high heat capacity is a property that hydrogen bonding among water molecules causes. Water has the highest **specific heat capacity** of any liquids. We define specific heat as the amount of heat one gram of a substance must absorb or lose to change its temperature by one degree Celsius. For water, this amount is one **calorie**. It therefore takes water a long time to heat and a long time to cool. In fact, water's specific heat capacity is about five times more than that of sand. This explains why the land cools faster than the sea. Due to its high heat capacity, warm blooded animals use water to more evenly disperse heat in their bodies: it acts in a similar manner to a car's cooling system, transporting heat from warm places to cool places, causing the body to maintain a more even temperature.

Water's Heat of Vaporization

Water also has a high **heat of vaporization**, the amount of energy required to change one gram of a liquid substance to a gas. A considerable amount of heat energy (586 cal) is required to accomplish this change in water. This process occurs on the water's surface. As liquid water heats up, hydrogen bonding makes it difficult to separate the liquid water molecules from each other, which is required for it to enter its gaseous phase (steam). As a result, water acts as a heat sink or heat reservoir and requires much more heat to boil than does a liquid such as ethanol (grain alcohol), whose hydrogen bonding with other ethanol molecules is weaker than water's hydrogen bonding. Eventually, as water reaches its boiling point of 100° Celsius (212° Fahrenheit), the heat is able to break the hydrogen bonds between the water molecules, and the kinetic energy (motion) between the water molecules allows them to escape from the liquid as a gas. Even when below its boiling point, water's individual molecules acquire enough energy from other water molecules such that some surface water molecules can escape and vaporize: we call this process **evaporation**.

The fact that hydrogen bonds need to be broken for water to evaporate means that bonds use a substantial amount of energy in the process. As the water evaporates, energy is taken up by the process, cooling the environment where the evaporation is taking place. In many living organisms, including in humans, the evaporation of sweat, which is 90

percent water, allows the organism to cool so that it can maintain homeostasis of body temperature.

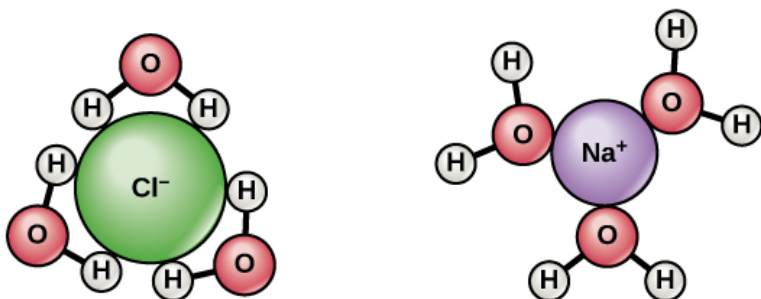
When we mix table salt (NaCl) in water, it forms spheres of hydration around the ions.

Water's Solvent Properties

Since water is a polar molecule with slightly positive and slightly negative charges, ions and polar molecules can readily dissolve in it. Therefore, we refer to water as a **solvent**, a substance capable of dissolving other polar molecules and ionic compounds. The charges associated with these molecules will form hydrogen bonds with water, surrounding the particle with water molecules. We refer to this as a **sphere of hydration**, or a hydration shell, as [\[link\]](#) illustrates and serves to keep the particles separated or dispersed in the water.

When we add ionic compounds to water, the individual ions react with the water molecules' polar regions and their ionic bonds are disrupted in the process of **dissociation**. Dissociation occurs when atoms or groups of atoms break off from molecules and form ions. Consider table salt (NaCl, or sodium chloride): when we add NaCl crystals to water, the NaCl molecules dissociate into Na^+ and Cl^- ions, and spheres of hydration form around the ions, as [\[link\]](#) illustrates. The partially negative charge of the water molecule's oxygen surrounds the

positively charged sodium ion. The hydrogen's partially positive charge on the water molecule surrounds the negatively charged chloride ion.



A needle's weight pulls the surface downward. At the same time, the surface tension pulls it up, suspending it on the water's surface preventing it from sinking. Notice the indentation in the water around the needle. (credit: Cory Zanker) The adhesive forces exerted by the glass' internal surface exceeding the cohesive forces between the water molecules themselves causes capillary action in a glass tube. (credit: modification of work by Pearson-Scott Foresman, donated to the Wikimedia Foundation) Water's cohesive and adhesive properties allow this water strider (*Gerris* sp.) to stay afloat. (credit: Tim Vickers)

Water's Cohesive and Adhesive Properties

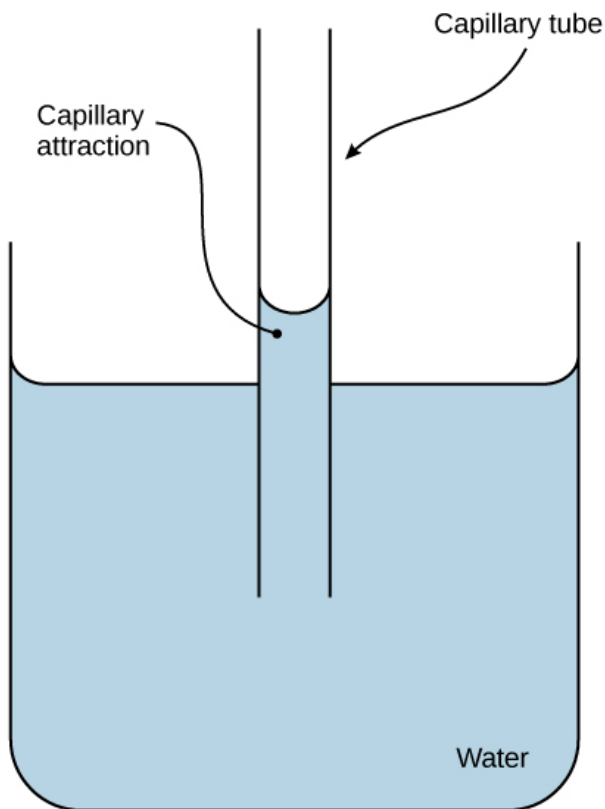
Have you ever filled a glass of water to the very top and then slowly added a few more drops? Before it overflows, the water forms a dome-like shape above the rim of the glass. This water can stay above the

glass because of the property of **cohesion**. In cohesion, water molecules are attracted to each other (because of hydrogen bonding), keeping the molecules together at the liquid-gas (water-air) interface, although there is no more room in the glass.

Cohesion allows for **surface tension**, the capacity of a substance to withstand rupturing when placed under tension or stress. This is also why water forms droplets when on a dry surface rather than flattening by gravity. When we place a small scrap of paper onto a water droplet, the paper floats on top even though paper is denser (heavier) than the water. Cohesion and surface tension keep the water molecules' hydrogen bonds intact and support the item floating on the top. It's even possible to “float” a needle on top of a glass of water if you place it gently without breaking the surface tension, as [\[link\]](#) shows.



These cohesive forces are related to water's property of **adhesion**, or the attraction between water molecules and other molecules. This attraction is sometimes stronger than water's cohesive forces, especially when the water is exposed to charged surfaces such as those on the inside of thin glass tubes known as capillary tubes. We observe adhesion when water "climbs" up the tube placed in a glass of water: notice that the water appears to be higher on the tube's sides than in the middle. This is because the water molecules are attracted to the capillary's charged glass walls more than they are to each other and therefore adhere to it. We call this type of adhesion **capillary action**, as [\[link\]](#) illustrates.



Why are cohesive and adhesive forces important for life? Cohesive and adhesive forces are important for transporting water from the roots to the leaves in plants. These forces create a “pull” on the water column. This pull results from the tendency of water molecules evaporating on the plant's surface to stay connected to water molecules below them, and so they are pulled along. Plants use this natural phenomenon to help transport water from their roots to their leaves. Without these properties of water, plants would be unable to receive the water and the dissolved minerals they require. In another

example, insects such as the water strider, as [\[link\]](#) shows, use the water's surface tension to stay afloat on the water's surface layer and even mate there.



The pH scale measures hydrogen ions' (H^+) concentration in a solution. (credit: modification of work by Edward Stevens) This diagram shows the body's buffering of blood pH levels. The blue arrows show the process of raising pH as more CO_2 is made. The purple arrows indicate the reverse process: the lowering of pH as more bicarbonate is created.

pH, Buffers, Acids, and Bases

The pH of a solution indicates its acidity or basicity.

$$H_2O(l) \leftrightarrow H^+(aq) + OH^-(aq)$$

You may have used **litmus** or pH paper, filter paper treated with a natural water-soluble dye for use as a pH indicator, tests how much acid (acidity) or base (basicity) exists in a solution. You might have even used some to test whether the water in a swimming

pool is properly treated. In both cases, the pH test measures hydrogen ions' concentration in a given solution.

Hydrogen ions spontaneously generate in pure water by the dissociation (ionization) of a small percentage of water molecules into equal numbers of hydrogen (H^+) ions and hydroxide (OH^-) ions. While the hydroxide ions are kept in solution by their hydrogen bonding with other water molecules, the hydrogen ions, consisting of naked protons, immediately attract to un-ionized water molecules, forming hydronium ions (H_3O^+). Still, by convention, scientists refer to hydrogen ions and their concentration as if they were free in this state in liquid water.

The concentration of hydrogen ions dissociating from pure water is 1×10^{-7} moles H^+ ions per liter of water. Moles (mol) are a way to express the amount of a substance (which can be atoms, molecules, ions, etc.). One mole represents the atomic weight of a substance, expressed in grams, which equals the amount of the substance containing as many units as there are atoms in 12 grams of ^{12}C . Mathematically, one mole is equal to 6.02×10^{23} particles of the substance. Therefore, 1 mole of water is equal to 6.02×10^{23} water molecules. We calculate the pH as the negative of the base 10 logarithm of this concentration. The \log_{10} of 1×10^{-7} is -7.0, and the negative of this

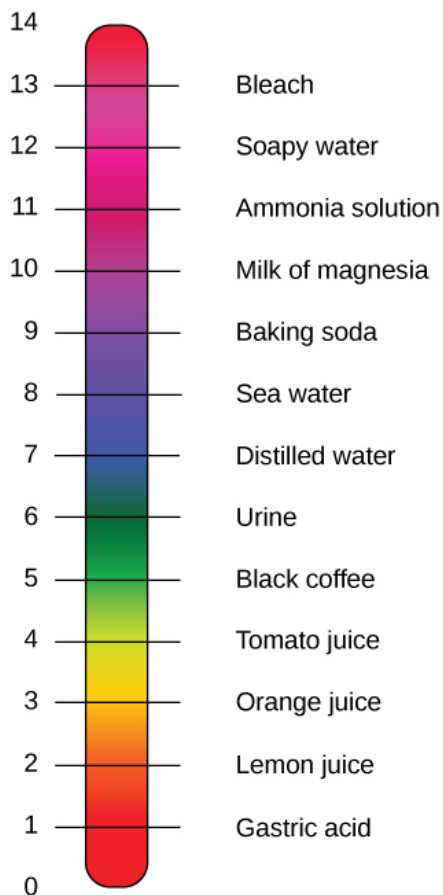
number (indicated by the “p” of “pH”) yields a pH of 7.0, which is also a neutral pH. The pH inside of human cells and blood are examples of two body areas where near-neutral pH is maintained.

Non-neutral pH readings result from dissolving acids or bases in water. Using the negative logarithm to generate positive integers, high concentrations of hydrogen ions yield a low pH number; whereas, low levels of hydrogen ions result in a high pH. An **acid** is a substance that increases hydrogen ions' (H^+) concentration in a solution, usually by having one of its hydrogen atoms dissociate. A **base** provides either hydroxide ions (OH^-) or other negatively charged ions that combine with hydrogen ions, reducing their concentration in the solution and thereby raising the pH. In cases where the base releases hydroxide ions, these ions bind to free hydrogen ions, generating new water molecules.

The stronger the acid, the more readily it donates H^+ . For example, hydrochloric acid (HCl) completely dissociates into hydrogen and chloride ions and is highly acidic; whereas the acids in tomato juice or vinegar do not completely dissociate and are weak acids. Conversely, strong bases are those substances that readily donate OH^- or take up hydrogen ions. Sodium hydroxide ($NaOH$) and many household cleaners are highly alkaline and give up OH^- rapidly when we place them in water, thereby raising the pH. An example of a weak basic solution is

seawater, which has a pH near 8.0 This is close enough to a neutral pH that marine organisms have adapted in order to live and thrive in a saline environment.

The **pH scale** is, as we previously mentioned, an inverse logarithm and ranges from 0 to 14 ([\[link\]](#)). Anything below 7.0 (ranging from 0.0 to 6.9) is acidic, and anything above 7.0 (from 7.1 to 14.0) is alkaline. Extremes in pH in either direction from 7.0 are usually inhospitable to life. The pH inside cells (6.8) and the pH in the blood (7.4) are both very close to neutral. However, the environment in the stomach is highly acidic, with a pH of 1 to 2. As a result, how do stomach cells survive in such an acidic environment? How do they homeostatically maintain the near neutral pH inside them? The answer is that they cannot do it and are constantly dying. The stomach constantly produces new cells to replace dead ones, which stomach acids digest. Scientists estimate that the human body completely replaces the stomach lining every seven to ten days.



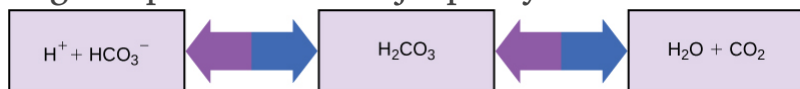
Link to Learning

Watch this video for a straightforward explanation of pH and its logarithmic scale.

https://www.openstax.org/l/pH_scale

How can organisms whose bodies require a near-

neutral pH ingest acidic and basic substances (a human drinking orange juice, for example) and survive? Buffers are the key. **Buffers** readily absorb excess H^+ or OH^- , keeping the body's pH carefully maintained in the narrow range required for survival. Maintaining a constant blood pH is critical to a person's well-being. The buffer maintaining the pH of human blood involves carbonic acid (H_2CO_3), bicarbonate ion (HCO_3^-), and carbon dioxide (CO_2). When bicarbonate ions combine with free hydrogen ions and become carbonic acid, it removes hydrogen ions and moderates pH changes. Similarly, as [\[link\]](#) shows, excess carbonic acid can convert to carbon dioxide gas which we exhale through the lungs. This prevents too many free hydrogen ions from building up in the blood and dangerously reducing the blood's pH. Likewise, if too much OH^- enters into the system, carbonic acid will combine with it to create bicarbonate, lowering the pH. Without this buffer system, the body's pH would fluctuate enough to put survival in jeopardy.



Other examples of buffers are antacids that some people use to combat excess stomach acid. Many of these over-the-counter medications work in the same way as blood buffers, usually with at least one ion capable of absorbing hydrogen and moderating pH, bringing relief to those who suffer “heartburn” after eating. Water's unique properties that

contribute to this capacity to balance pH—as well as water’s other characteristics—are essential to sustaining life on Earth.

Link to Learning

To learn more about water, visit the [U.S. Geological Survey Water Science for Schools All About Water!](#) website.

Section Summary

Water has many properties that are critical to maintaining life. It is a polar molecule, allowing for forming hydrogen bonds. Hydrogen bonds allow ions and other polar molecules to dissolve in water. Therefore, water is an excellent solvent. The hydrogen bonds between water molecules cause the water to have a high heat capacity, meaning it takes considerable added heat to raise its temperature. As the temperature rises, the hydrogen bonds between water continually break and form anew. This allows for the overall temperature to remain stable, although energy is added to the system. Water also exhibits a high heat of vaporization, which is key to how organisms cool themselves by evaporating

sweat. Water's cohesive forces allow for the property of surface tension; whereas, we see its adhesive properties as water rises inside capillary tubes. The pH value is a measure of hydrogen ion concentration in a solution and is one of many chemical characteristics that is highly regulated in living organisms through homeostasis. Acids and bases can change pH values, but buffers tend to moderate the changes they cause. These properties of water are intimately connected to the biochemical and physical processes performed by living organisms, and life would be very different if these properties were altered, if it could exist at all.

Review Questions

Which of the following statements is not true?

1. Water is polar.
2. Water stabilizes temperature.
3. Water is essential for life.
4. Water is the most abundant molecule in the Earth's atmosphere.

D

When acids are added to a solution, the pH

should _____.

1. decrease
 2. increase
 3. stay the same
 4. cannot tell without testing
-

A

We call a molecule that binds up excess hydrogen ions in a solution a(n) _____.

1. acid
 2. isotope
 3. base
 4. donator
-

C

Which of the following statements is true?

1. Acids and bases cannot mix together.
 2. Acids and bases will neutralize each other.
 3. Acids, but not bases, can change the pH of a solution.
 4. Acids donate hydroxide ions (OH^-); bases donate hydrogen ions (H^+).
-

Critical Thinking Questions

Discuss how buffers help prevent drastic swings in pH.

Buffers absorb the free hydrogen ions and hydroxide ions that result from chemical reactions. Because they can bond these ions, they prevent increases or decreases in pH. An example of a buffer system is the bicarbonate system in the human body. This system is able to absorb hydrogen and hydroxide ions to prevent changes in pH and keep cells functioning properly.

Why can some insects walk on water?

Some insects can walk on water, although they are heavier (denser) than water, because of the surface tension of water. Surface tension results from cohesion, or the attraction between water molecules at the surface of the body of water (the liquid-air/gas interface).

Glossary

acid

molecule that donates hydrogen ions and increases the concentration of hydrogen ions in a solution

adhesion

attraction between water molecules and other molecules

base

molecule that donates hydroxide ions or otherwise binds excess hydrogen ions and decreases the hydrogen ions' concentration in a solution

buffer

substance that prevents a change in pH by absorbing or releasing hydrogen or hydroxide ions

calorie

amount of heat required to change the temperature of one gram of water by one degree Celsius

capillary action

occurs because water molecules are attracted to charges on the inner surfaces of narrow tubular structures such as glass tubes, drawing the water molecules to the tubes'

sides

cohesion

intermolecular forces between water molecules caused by the polar nature of water; responsible for surface tension

dissociation

release of an ion from a molecule such that the original molecule now consists of an ion and the charged remains of the original, such as when water dissociates into H^+ and OH^-

evaporation

change from liquid to gaseous state at a body of water's surface, plant leaves, or an organism's skin

heat of vaporization of water

high amount of energy required for liquid water to turn into water vapor

hydrophilic

describes ions or polar molecules that interact well with other polar molecules such as water

hydrophobic

describes uncharged nonpolar molecules that do not interact well with polar molecules such as water

litmus paper

(also, pH paper) filter paper treated with a natural water-soluble dye that changes its color as the pH of the environment changes in order to use it as a pH indicator

pH paper

see litmus paper

pH scale

scale ranging from zero to 14 that is inversely proportional to the hydrogen ions' concentration in a solution

solvent

substance capable of dissolving another substance

specific heat capacity

the amount of heat one gram of a substance must absorb or lose to change its temperature by one degree Celsius

sphere of hydration

when a polar water molecule surrounds charged or polar molecules thus keeping them dissolved and in solution

surface tension

tension at the surface of a body of liquid that prevents the molecules from separating; created by the attractive cohesive forces between the liquid's molecules

Carbon

By the end of this section, you will be able to do the following:

- Explain why carbon is important for life
- Describe the role of functional groups in biological molecules

Many complex molecules called macromolecules, such as proteins, nucleic acids (RNA and DNA), carbohydrates, and lipids comprise cells. The macromolecules are a subset of **organic molecules** (any carbon-containing liquid, solid, or gas) that are especially important for life. The fundamental component for all of these macromolecules is carbon. The carbon atom has unique properties that allow it to form covalent bonds to as many as four different atoms, making this versatile element ideal to serve as the basic structural component, or “backbone,” of the macromolecules.

Individual carbon atoms have an incomplete outermost electron shell. With an atomic number of 6 (six electrons and six protons), the first two electrons fill the inner shell, leaving four in the second shell. Therefore, carbon atoms can form up to four covalent bonds with other atoms to satisfy the octet rule. The methane molecule provides an example: it has the chemical formula CH_4 . Each of its four hydrogen atoms forms a single covalent bond with the carbon atom by sharing a pair of

electrons. This results in a filled outermost shell.

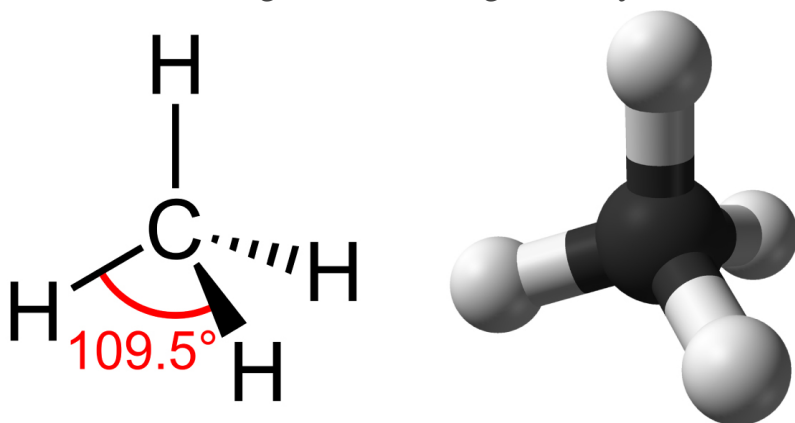
Methane has a tetrahedral geometry, with each of the four hydrogen atoms spaced 109.5° apart. When carbon forms single bonds with other atoms, the shape is tetrahedral. When two carbon atoms form a double bond, the shape is planar, or flat. Single bonds, like those in ethane, are able to rotate.

Double bonds, like those in ethene, cannot rotate, so the atoms on either side are locked in place. Carbon can form five- and six-membered rings. Single or double bonds may connect the carbons in the ring, and nitrogen may be substituted for carbon.

Hydrocarbons

Hydrocarbons are organic molecules consisting entirely of carbon and hydrogen, such as methane (CH_4) described above. We often use hydrocarbons in our daily lives as fuels—like the propane in a gas grill or the butane in a lighter. The many covalent bonds between the atoms in hydrocarbons store a great amount of energy, which releases when these molecules burn (oxidize). Methane, an excellent fuel, is the simplest hydrocarbon molecule, with a central carbon atom bonded to four different hydrogen atoms, as [\[link\]](#) illustrates. The shape of its electron orbitals determines the shape of the methane molecule's geometry, where the atoms reside in three dimensions. The carbons and the four hydrogen atoms form a tetrahedron, with four triangular faces. For this reason, we describe

methane as having tetrahedral geometry.

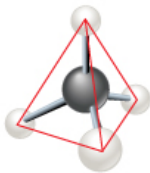
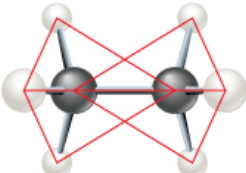
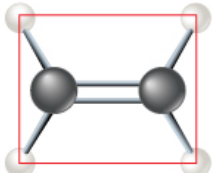


As the backbone of the large molecules of living things, hydrocarbons may exist as linear carbon chains, carbon rings, or combinations of both. Furthermore, individual carbon-to-carbon bonds may be single, double, or triple covalent bonds, and each type of bond affects the molecule's geometry in a specific way. This three-dimensional shape or conformation of the large molecules of life (macromolecules) is critical to how they function.

Hydrocarbon Chains

Successive bonds between carbon atoms form hydrocarbon chains. These may be branched or unbranched. Furthermore, a molecule's different geometries of single, double, and triple covalent bonds alter the overall molecule's geometry as [\[link\]](#) illustrates. The hydrocarbons ethane, ethene, and ethyne serve as examples of how different carbon-

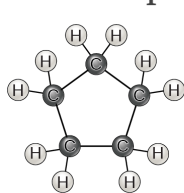
to-carbon bonds affect the molecule's geometry. The names of all three molecules start with the prefix “eth-,” which is the prefix for two carbon hydrocarbons. The suffixes “-ane,” “-ene,” and “-yne” refer to the presence of single, double, or triple carbon-carbon bonds, respectively. Thus, propane, propene, and propyne follow the same pattern with three carbon molecules, butane, butene, and butyne for four carbon molecules, and so on. Double and triple bonds change the molecule's geometry: single bonds allow rotation along the bond's axis; whereas, double bonds lead to a planar configuration and triple bonds to a linear one. These geometries have a significant impact on the shape a particular molecule can assume.

Methane (CH ₄)	Ethane (C ₂ H ₆)	Ethene (C ₂ H ₄)
		
Tetrahedral (single bond)	Tetrahedral (single bond)	Planar (double bond)

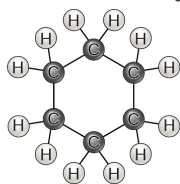
Hydrocarbon Rings

So far, the hydrocarbons we have discussed have been **aliphatic hydrocarbons**, which consist of linear chains of carbon atoms. Another type of hydrocarbon, **aromatic hydrocarbons**, consists of closed rings of carbon atoms. We find ring structures in hydrocarbons, sometimes with the presence of

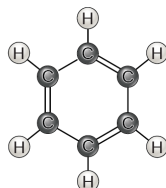
double bonds, which we can see by comparing cyclohexane's structure to benzene in [\[link\]](#). Examples of biological molecules that incorporate the benzene ring include some amino acids and cholesterol and its derivatives, including the hormones estrogen and testosterone. We also find the benzene ring in the herbicide 2,4-D. Benzene is a natural component of crude oil and has been classified as a carcinogen. Some hydrocarbons have both aliphatic and aromatic portions. Beta-carotene is an example of such a hydrocarbon.



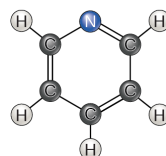
Cyclopentane



Cyclohexane



Benzene



Pyridine

These space-filling models show a *cis* (oleic acid) and a *trans* (eliadic acid) fatty acid. Notice the bend in the molecule caused by the *cis* configuration.

Isomers

The three-dimensional placement of atoms and chemical bonds within organic molecules is central to understanding their chemistry. We call molecules that share the same chemical formula but differ in the placement (structure) of their atoms and/or chemical bonds **isomers**. **Structural isomers** (like butane and isobutene in [\[link\]](#)[a](#)) differ in the placement of their covalent bonds: both molecules have four carbons and ten hydrogens (C_4H_{10}), but

the different atom arrangement within the molecules leads to differences in their chemical properties. For example, butane is suited for use as a fuel for cigarette lighters and torches; whereas, isobutene is suited for use as a refrigerant and a propellant in spray cans.

Geometric isomers, alternatively have similar placements of their covalent bonds but differ in how these bonds are made to the surrounding atoms, especially in carbon-to-carbon double bonds. In the simple molecule butene (C_4H_8), the two methyl groups (CH_3) can be on either side of the double covalent bond central to the molecule, as [\[link\]](#) illustrates. When the carbons are bound on the same side of the double bond, this is the *cis* configuration. If they are on opposite sides of the double bond, it is a *trans* configuration. In the *trans* configuration, the carbons form a more or less linear structure; whereas, the carbons in the *cis* configuration make a bend (change in direction) of the carbon backbone.

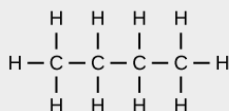
Visual Connection

We call molecules that have the same number and type of atoms arranged differently isomers. (a) Structural isomers have a different covalent arrangement of atoms. (b) Geometric isomers have a different arrangement of atoms around a double bond. (c) Enantiomers are mirror images of each

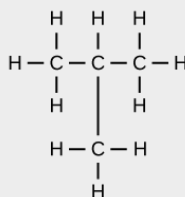
other.

(a) Structural isomers

Butane

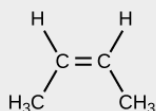


Isobutane



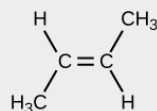
(b) Geometric isomers

cis-2-butene



methyl groups on
same side of double bond

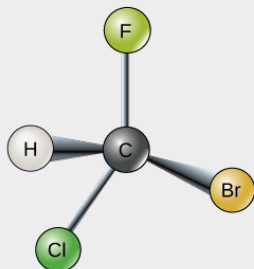
trans-2-butene



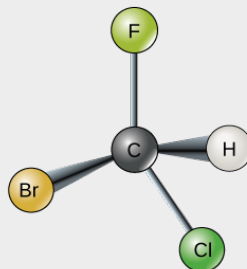
methyl groups on opposite
sides of double bond

(c) Enantiomers

L-isomer



D-isomer



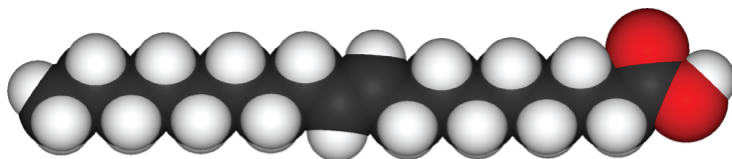
Which of the following statements is false?

1. Molecules with the formulas $\text{CH}_3\text{CH}_2\text{COOH}$ and $\text{C}_3\text{H}_6\text{O}_2$ could be structural isomers.
2. Molecules must have a double bond to be *cis-trans* isomers.
3. To be enantiomers, a molecule must have at

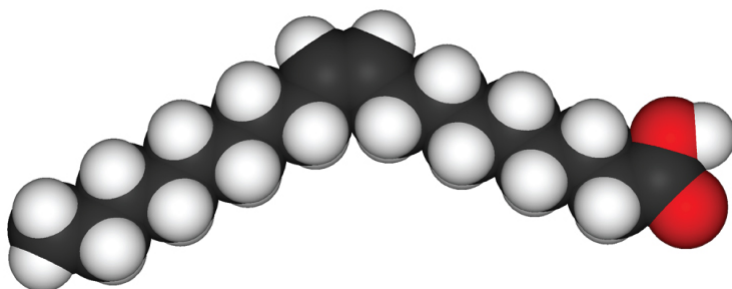
least three different atoms or groups connected to a central carbon.

4. To be enantiomers, a molecule must have at least four different atoms or groups connected to a central carbon.

In triglycerides (fats and oils), long carbon chains known as fatty acids may contain double bonds, which can be in either the *cis* or *trans* configuration, as [\[link\]](#) illustrates. Fats with at least one double bond between carbon atoms are unsaturated fats. When some of these bonds are in the *cis* configuration, the resulting bend in the chain's carbon backbone means that triglyceride molecules cannot pack tightly, so they remain liquid (oil) at room temperature. Alternatively, triglycerides with *trans* double bonds (popularly called trans fats), have relatively linear fatty acids that are able to pack tightly together at room temperature and form solid fats. In the human diet, trans fats are linked to an increased risk of cardiovascular disease, so many food manufacturers have reduced or eliminated their use in recent years. In contrast to unsaturated fats, we call triglycerides without double bonds between carbon atoms saturated fats, meaning that they contain all the hydrogen atoms available. Saturated fats are a solid at room temperature and usually of animal origin.



Elaidic acid



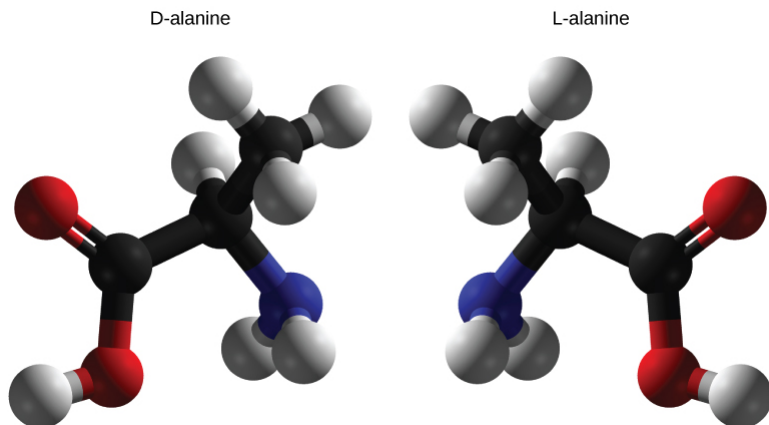
Oleic acid

D-alanine and L-alanine are examples of enantiomers or mirror images. L-forms of amino acids are predominant in proteins.

Enantiomers

Enantiomers are molecules that share the same chemical structure and chemical bonds but differ in the three-dimensional placement of atoms so that they are non-superimposable mirror images. [\[link\]](#) shows an amino acid alanine example, where the two structures are nonsuperimposable. In nature, the L-forms of amino acids are predominant in proteins. Some D forms of amino acids are seen in the cell walls of bacteria and polypeptides in other organisms. Similarly, the D-form of glucose is the main product of photosynthesis and we rarely see

the molecule's L-form in nature.



These functional groups are in many different biological molecules. R, also known as R-group, is an abbreviation for any group in which a carbon or hydrogen atom is attached to the rest of the molecule. Hydrogen bonds connect two strands of DNA together to create the double-helix structure.

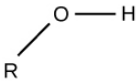
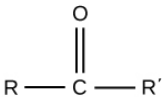
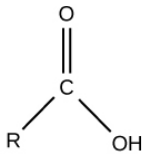
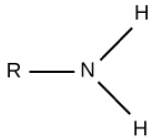
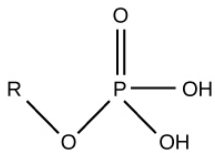
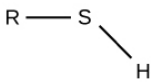
Functional Groups

Functional groups are groups of atoms that occur within molecules and confer specific chemical properties to those molecules. We find them along the “carbon backbone” of macromolecules. Chains and/or rings of carbon atoms with the occasional substitution of an element such as nitrogen or oxygen form this carbon backbone. Molecules with other elements in their carbon backbone are **substituted hydrocarbons**.

The functional groups in a macromolecule are

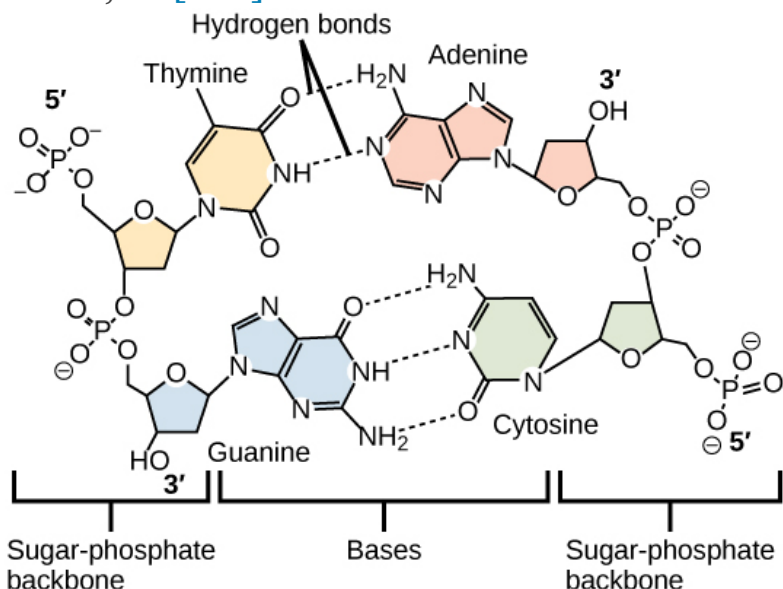
usually attached to the carbon backbone at one or several different places along its chain and/or ring structure. Each of the four types of macromolecules—proteins, lipids, carbohydrates, and nucleic acids—has its own characteristic set of functional groups that contributes greatly to its differing chemical properties and its function in living organisms.

A functional group can participate in specific chemical reactions. [\[link\]](#) shows some of the important functional groups in biological molecules. They include: hydroxyl, methyl, carbonyl, carboxyl, amino, phosphate, and sulfhydryl. These groups play an important role in forming molecules like DNA, proteins, carbohydrates, and lipids. We usually classify functional groups as hydrophobic or hydrophilic depending on their charge or polarity characteristics. An example of a hydrophobic group is the nonpolar methyl molecule. Among the hydrophilic functional groups is the carboxyl group in amino acids, some amino acid side chains, and the fatty acids that form triglycerides and phospholipids. This carboxyl group ionizes to release hydrogen ions (H^+) from the $COOH$ group resulting in the negatively charged COO^- group. This contributes to the hydrophilic nature of whatever molecule on which it is found. Other functional groups, such as the carbonyl group, have a partially negatively charged oxygen atom that may form hydrogen bonds with water molecules, again making the molecule more hydrophilic.

Functional Group	Structure	Properties
Hydroxyl		Polar
Methyl	$R - CH_3$	Nonpolar
Carbonyl		Polar
Carboxyl		Charged, ionizes to release H^+ . Since carboxyl groups can release H^+ ions into solution, they are considered acidic.
Amino		Charged, accepts H^+ to form NH_3^+ . Since amino groups can remove H^+ from solution, they are considered basic.
Phosphate		Charged, ionizes to release H^+ . Since phosphate groups can release H^+ ions into solution, they are considered acidic.
Sulfhydryl		Polar

Hydrogen bonds between functional groups (within the same molecule or between different molecules) are important to the function of many macromolecules and help them to fold properly into and maintain the appropriate shape for functioning. Hydrogen bonds are also involved in various recognition processes, such as DNA complementary base pairing and the binding of an enzyme to its

substrate, as [\[link\]](#) illustrates.



Section Summary

The unique properties of carbon make it a central part of biological molecules. Carbon binds to oxygen, hydrogen, and nitrogen covalently to form the many molecules important for cellular function. Carbon has four electrons in its outermost shell and can form four bonds. Carbon and hydrogen can form hydrocarbon chains or rings. Functional groups are groups of atoms that confer specific properties to hydrocarbon (or substituted hydrocarbon) chains or rings that define their overall chemical characteristics and function.

Visual Connection Questions

[\[link\]](#) Which of the following statements is false?

1. Molecules with the formulas $\text{CH}_3\text{CH}_2\text{COOH}$ and $\text{C}_3\text{H}_6\text{O}_2$ could be structural isomers.
2. Molecules must have a double bond to be *cis-trans* isomers.
3. To be enantiomers, a molecule must have at least three different atoms or groups connected to a central carbon.
4. To be enantiomers, a molecule must have at least four different atoms or groups connected to a central carbon.

[\[link\]](#) C

Review Questions

Each carbon molecule can bond with as many as _____ other atom(s) or molecule(s).

1. one
2. two

3. six
4. four

D

Which of the following is not a functional group that can bond with carbon?

1. sodium
2. hydroxyl
3. phosphate
4. carbonyl

A

Critical Thinking Questions

What property of carbon makes it essential for organic life?

Carbon is unique and found in all living things because it can form up to four covalent bonds between atoms or molecules. These can be nonpolar or polar covalent bonds, and they allow for the formation of long chains of carbon

molecules that combine to form proteins and DNA.

Compare and contrast saturated and unsaturated triglycerides.

Saturated triglycerides contain no double bonds between carbon atoms; they are usually solid at room temperature. Unsaturated triglycerides contain at least one double bond between carbon atoms and are usually liquid at room temperature.

Glossary

aliphatic hydrocarbon

hydrocarbon consisting of a linear chain of carbon atoms

aromatic hydrocarbon

hydrocarbon consisting of closed rings of carbon atoms

enantiomers

molecules that share overall structure and bonding patterns, but differ in how the atoms are three dimensionally placed such that they are mirror images of each other

functional group

group of atoms that provides or imparts a specific function to a carbon skeleton

geometric isomer

isomer with similar bonding patterns differing in the placement of atoms alongside a double covalent bond

hydrocarbon

molecule that consists only of carbon and hydrogen

isomers

molecules that differ from one another even though they share the same chemical formula

organic molecule

any molecule containing carbon (except carbon dioxide)

structural isomers

molecules that share a chemical formula but differ in the placement of their chemical bonds

substituted hydrocarbon

hydrocarbon chain or ring containing an atom of another element in place of one of the backbone carbons

Introduction

class = "introduction" Foods such as bread, fruit, and cheese are rich sources of biological macromolecules. (credit: modification of work by Bengt Nyman)



Food provides the body with the nutrients it needs to survive. Many of these critical nutrients are biological macromolecules, or large molecules, necessary for life. Different smaller organic molecule (monomer) combinations build these macromolecules (polymers). What specific biological macromolecules do living things require? How do these molecules form? What functions do they serve? We explore these questions in this chapter.

Synthesis of Biological Macromolecules

By the end of this section, you will be able to do the following:

- Understand macromolecule synthesis
- Explain dehydration (or condensation) and hydrolysis reactions

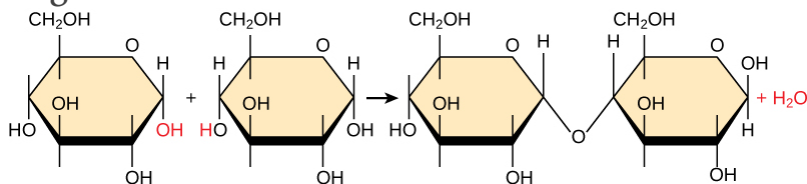
As you've learned, **biological macromolecules** are large molecules, necessary for life, that are built from smaller organic molecules. There are four major biological macromolecule classes (carbohydrates, lipids, proteins, and nucleic acids). Each is an important cell component and performs a wide array of functions. Combined, these molecules make up the majority of a cell's dry mass (recall that water makes up the majority of its complete mass). Biological macromolecules are organic, meaning they contain carbon. In addition, they may contain hydrogen, oxygen, nitrogen, and additional minor elements.

In the dehydration synthesis reaction above, two glucose molecules link to form the disaccharide maltose. In the process, it forms a water molecule.

Dehydration Synthesis

Most macromolecules are made from single subunits, or building blocks, called **monomers**. The monomers combine with each other using covalent

bonds to form larger molecules known as **polymers**. In doing so, monomers release water molecules as byproducts. This type of reaction is **dehydration synthesis**, which means “to put together while losing water.”



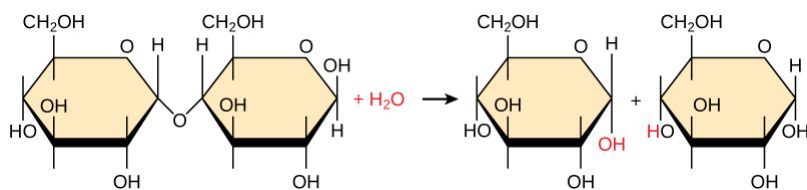
In a dehydration synthesis reaction ([\[link\]](#)), the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a water molecule. At the same time, the monomers share electrons and form covalent bonds. As additional monomers join, this chain of repeating monomers forms a polymer. Different monomer types can combine in many configurations, giving rise to a diverse group of macromolecules. Even one kind of monomer can combine in a variety of ways to form several different polymers. For example, glucose monomers are the constituents of starch, glycogen, and cellulose.

In the hydrolysis reaction here, the disaccharide maltose breaks down to form two glucose monomers by adding a water molecule. Note that this reaction is the reverse of the synthesis reaction in [\[link\]](#).

Hydrolysis

Polymers break down into monomers during

hydrolysis. A chemical reaction occurs when inserting a water molecule across the bond. Breaking a covalent bond with this water molecule in the compound achieves this ([\[link\]](#)). During these reactions, the polymer breaks into two components: one part gains a hydrogen atom (H^+) and the other gains a hydroxyl molecule (OH^-) from a split water molecule.



Dehydration and **hydrolysis reactions** are catalyzed, or “sped up,” by specific enzymes; dehydration reactions involve the formation of new bonds, requiring energy, while hydrolysis reactions break bonds and release energy. These reactions are similar for most macromolecules, but each monomer and polymer reaction is specific for its class. For example, catalytic enzymes in the digestive system hydrolyze or break down the food we ingest into smaller molecules. This allows cells in our body to easily absorb nutrients in the intestine. A specific enzyme breaks down each macromolecule. For instance, amylase, sucrase, lactase, or maltase break down carbohydrates. Enzymes called proteases, such as pepsin and peptidase, and hydrochloric acid break down proteins. Lipases break down lipids. These broken down macromolecules provide energy for cellular activities.

Link to Learning

Visit [this site](#) to see visual representations of dehydration synthesis and hydrolysis.

Section Summary

Proteins, carbohydrates, nucleic acids, and lipids are the four major classes of biological macromolecules—large molecules necessary for life that are built from smaller organic molecules. Macromolecules are comprised of single units scientists call monomers that are joined by covalent bonds to form larger polymers. The polymer is more than the sum of its parts: it acquires new characteristics, and leads to an osmotic pressure that is much lower than that formed by its ingredients. This is an important advantage in maintaining cellular osmotic conditions. A monomer joins with another monomer with water molecule release, leading to a covalent bond forming. Scientists call these dehydration or condensation reactions. When polymers break down into smaller units (monomers), they use a water molecule for each bond broken by these reactions. Such reactions are hydrolysis reactions. Dehydration and hydrolysis reactions are similar for all macromolecules, but each monomer and polymer

reaction is specific to its class. Dehydration reactions typically require an investment of energy for new bond formation, while hydrolysis reactions typically release energy by breaking bonds.

Review Questions

Dehydration synthesis leads to formation of

1. monomers
2. polymers
3. water and polymers
4. none of the above

C

During the breakdown of polymers, which of the following reactions takes place?

1. hydrolysis
2. dehydration
3. condensation
4. covalent bond

A

The following chemical reactants produce the ester ethyl ethanoate ($\text{C}_4\text{H}_8\text{O}_2$):



What type of reaction occurs to make ethyl ethanoate?

1. condensation
2. hydrolysis
3. combustion
4. acid-base reaction

A

Critical Thinking Questions

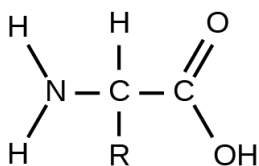
Why are biological macromolecules considered organic?

Biological macromolecules are organic because they contain carbon.

What role do electrons play in dehydration synthesis and hydrolysis?

In a dehydration synthesis reaction, the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a molecule of water. This creates an opening in the outer shells of atoms in the monomers, which can share electrons and form covalent bonds.

Amino acids have the generic structure seen below, where R represents different carbon-based side chains.



Describe how the structure of amino acids allows them to be linked into long peptide chains to form proteins.

Amino acids can be linked into long chains through condensation reactions. One of the hydrogen atoms bonded to the nitrogen atom of an amino acid reacts with the -OH group attached to the terminal carbon on another amino acid. Since both ends of the molecule can participate in condensation reactions, peptide bonds can be made in both directions to

create a long amino acid chain.

Glossary

biological macromolecule

large molecule necessary for life that is built from smaller organic molecules

dehydration synthesis

(also, condensation) reaction that links monomer molecules, releasing a water molecule for each bond formed

hydrolysis

reaction that causes breakdown of larger molecules into smaller molecules by utilizing water

monomer

smallest unit of larger molecules that are polymers

polymer

chain of monomer residues that covalent bonds link; polymerization is the process of polymer formation from monomers by condensation

Carbohydrates

By the end of this section, you will be able to do the following:

- Discuss the role of carbohydrates in cells and in the extracellular materials of animals and plants
- Explain carbohydrate classifications
- List common monosaccharides, disaccharides, and polysaccharides

Most people are familiar with carbohydrates, one type of macromolecule, especially when it comes to what we eat. To lose weight, some individuals adhere to “low-carb” diets. Athletes, in contrast, often “carb-load” before important competitions to ensure that they have enough energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet. Grains, fruits, and vegetables are all natural carbohydrate sources that provide energy to the body, particularly through glucose, a simple sugar that is a component of **starch** and an ingredient in many staple foods. Carbohydrates also have other important functions in humans, animals, and plants.

Scientists classify monosaccharides based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain, and ketoses have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and

hexoses have three-, five-, and six- carbon backbones, respectively. Five and six carbon monosaccharides exist in equilibrium between linear and ring forms. When the ring forms, the side chain it closes on locks into an α or β position. Fructose and ribose also form rings, although they form five-membered rings as opposed to the six-membered ring of glucose. Sucrose forms when a glucose monomer and a fructose monomer join in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage forms between carbon 1 in glucose and carbon 2 in fructose. Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar). Amylose and amylopectin are two different starch forms. Unbranched glucose monomer chains comprise amylose by α 1-4 glycosidic linkages. Unbranched glucose monomer chains comprise amylopectin by α 1-4 and α 1-6 glycosidic linkages. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched. In cellulose, glucose monomers are linked in unbranched chains by β 1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)

Molecular Structures

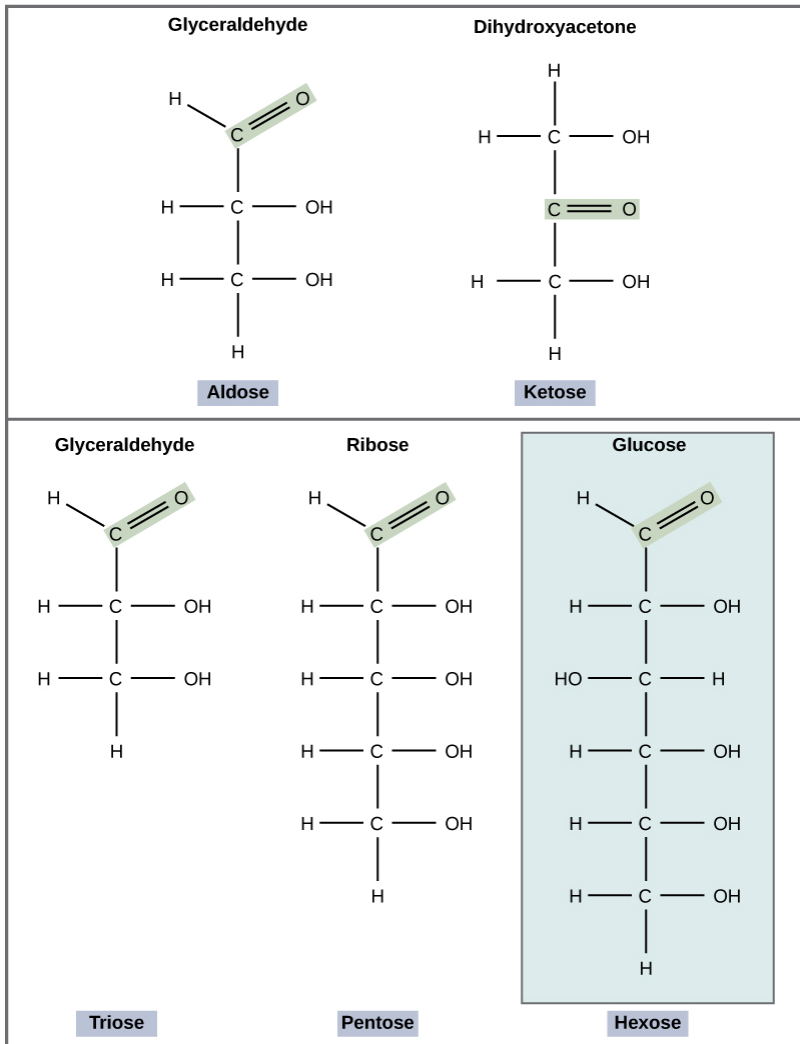
The stoichiometric formula $(\text{CH}_2\text{O})_n$, where n is the number of carbons in the molecule represents **carbohydrates**. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term “carbohydrate”: the components are carbon (“carbo”) and the components of water (hence, “hydrate”). Scientists classify carbohydrates into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides

Monosaccharides (mono- = “one”; sacchar- = “sweet”) are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. Most monosaccharide names end with the suffix -ose. If the sugar has an aldehyde group (the functional group with the structure R-CHO), it is an aldose, and if it has a ketone group (the functional group with the structure RC(=O)R'), it is a ketose. Depending on the number of carbons in the sugar, they can be trioses (three carbons), pentoses (five

carbons), and/or hexoses (six carbons). [\[link\]](#) illustrates monosaccharides.

MONOSACCHARIDES



The chemical formula for glucose is $C_6H_{12}O_6$. In humans, glucose is an important source of energy. During cellular respiration, energy releases from glucose, and that energy helps make adenosine

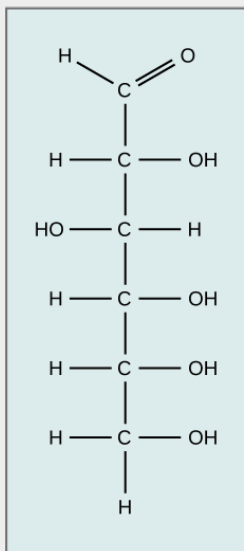
triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn provides energy requirements for the plant. Humans and other animals that feed on plants often store excess glucose as catabolized (cell breakdown of larger molecules) starch.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are isomers) because of the different arrangement of functional groups around the asymmetric carbon. All these monosaccharides have more than one asymmetric carbon ([\[link\]](#)).

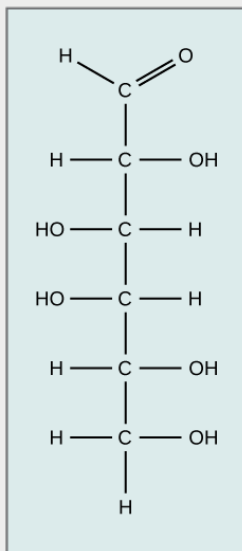
Visual Connection

Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different atom arrangement.

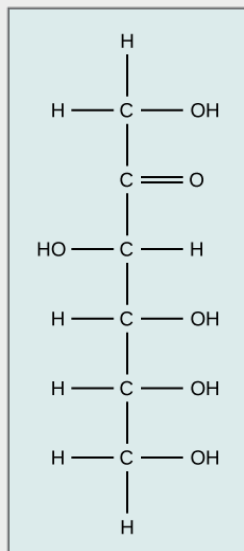
Glucose



Galactose



Fructose

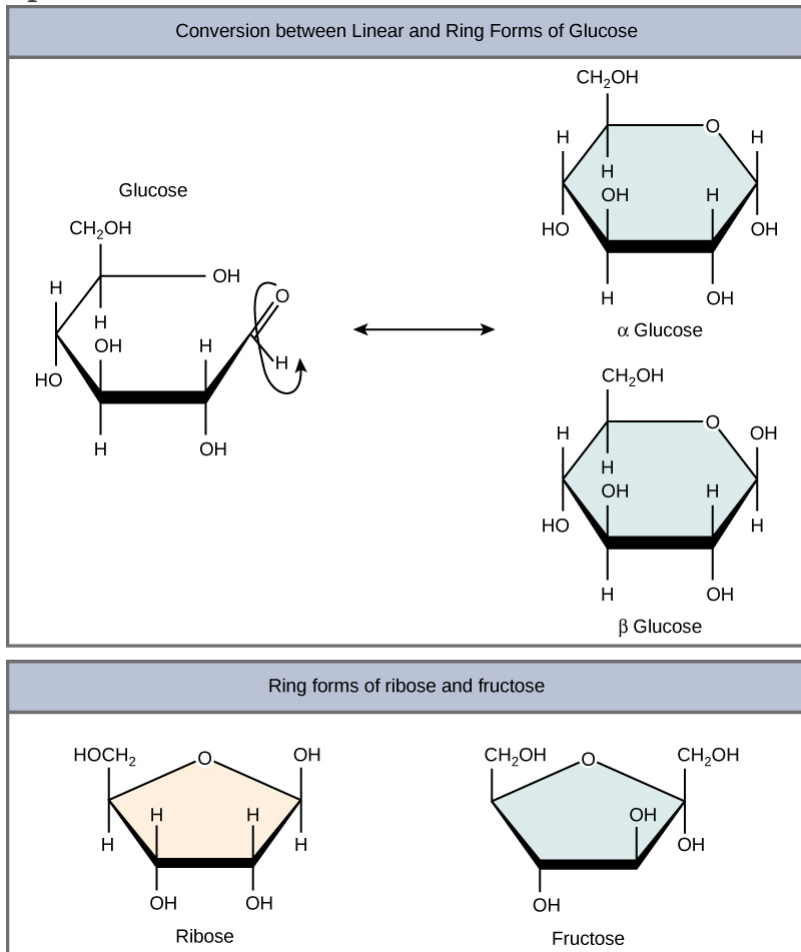


What kind of sugars are these, aldose or ketose?

Glucose, galactose, and fructose are isomeric monosaccharides (hexoses), meaning they have the same chemical formula but have slightly different structures. Glucose and galactose are aldoses, and fructose is a ketose.

Monosaccharides can exist as a linear chain or as ring-shaped molecules. In aqueous solutions they are usually in ring forms ([\[link\]](#)). Glucose in a ring form can have two different hydroxyl group arrangements (OH) around the anomeric carbon (carbon 1 that becomes asymmetric in the ring formation process). If the hydroxyl group is below carbon number 1 in the sugar, it is in the alpha (α)

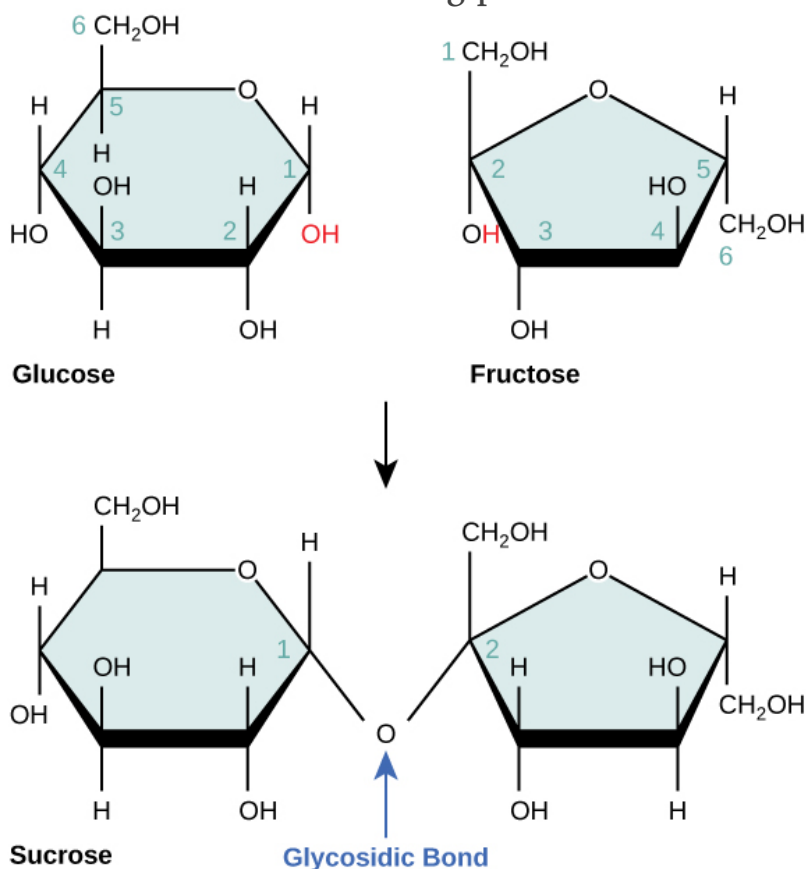
position, and if it is above the plane, it is in the beta (β) position.



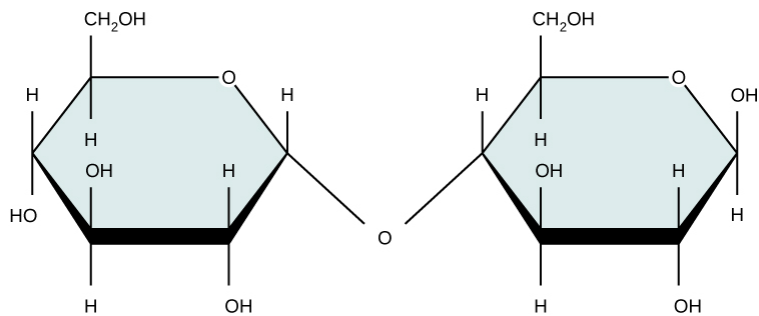
Disaccharides

Disaccharides (di- = “two”) form when two monosaccharides undergo a dehydration reaction (or a condensation reaction or dehydration synthesis). During this process, one monosaccharide's hydroxyl group combines with

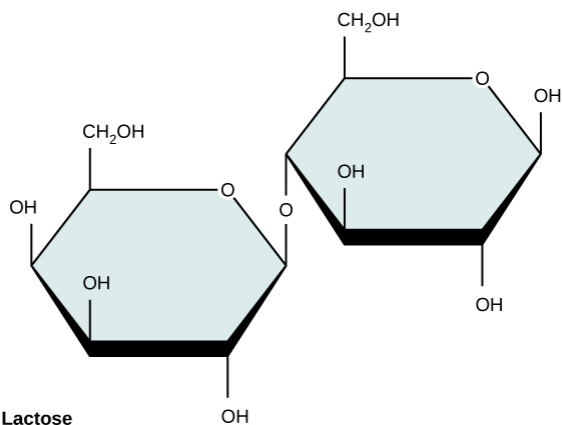
another monosaccharide's hydrogen, releasing a water molecule and forming a covalent bond. A covalent bond forms between a carbohydrate molecule and another molecule (in this case, between two monosaccharides). Scientists call this a **glycosidic bond** ([\[link\]](#)). Glycosidic bonds (or glycosidic linkages) can be an alpha or beta type. An alpha bond is formed when the OH group on the carbon-1 of the first glucose is below the ring plane, and a beta bond is formed when the OH group on the carbon-1 is above the ring plane.



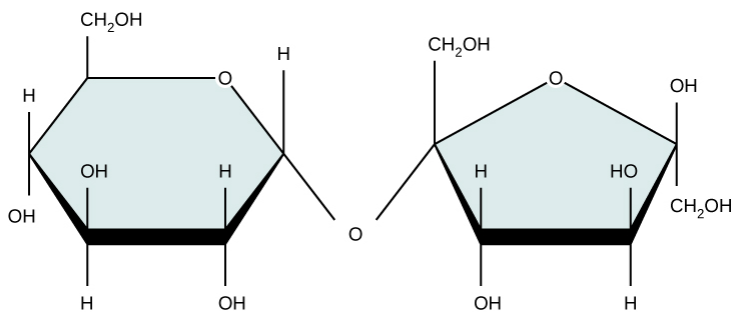
Common disaccharides include lactose, maltose, and sucrose ([\[link\]](#)). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is naturally in milk. Maltose, or malt sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is comprised of glucose and fructose monomers.



Maltose



Lactose



Sucrose

Polysaccharides

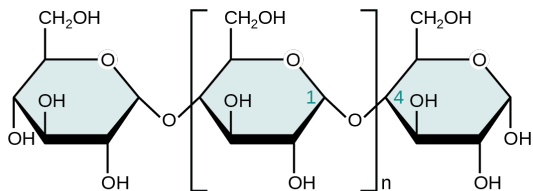
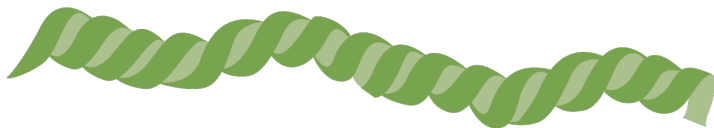
A long chain of monosaccharides linked by glycosidic bonds is a **polysaccharide** (poly- =

“many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 daltons or more depending on the number of joined monomers. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

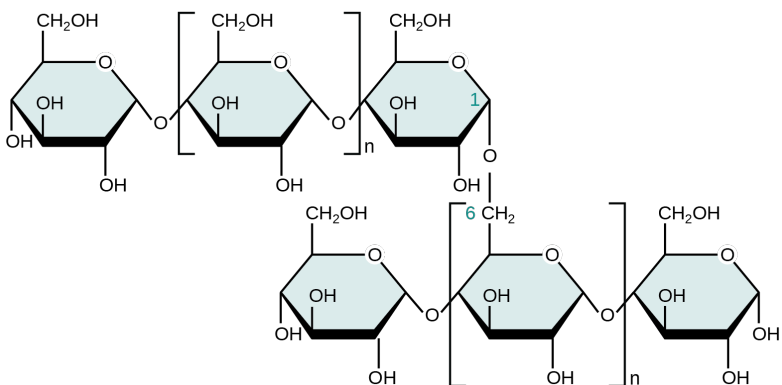
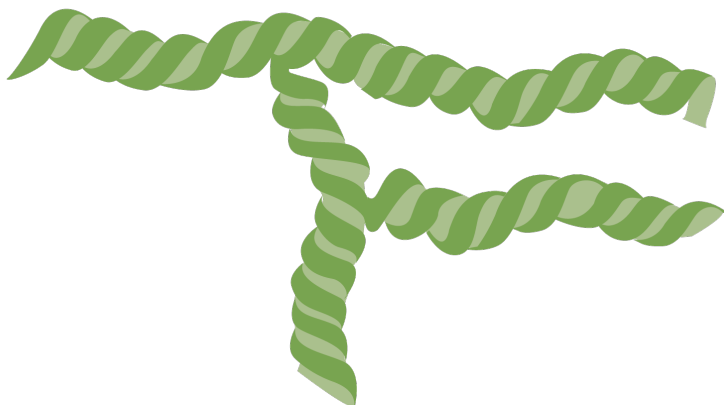
Plants store starch in the form of sugars. In plants, an amylose and amylopectin mixture (both glucose polymers) comprise these sugars. Plants are able to synthesize glucose, and they store the excess glucose, beyond their immediate energy needs, as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a food source for humans and animals. Enzymes break down the starch that humans consume. For example, an amylase present in saliva catalyzes, or breaks down this starch into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

Glucose starch comprises monomers that are joined by α 1-4 or α 1-6 glycosidic bonds. The numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As [\[link\]](#) illustrates, unbranched glucose monomer chains (only α 1-4 linkages) form the starch; whereas, amylopectin is a branched polysaccharide (α 1-6 linkages at the branch points).

Amylose



Amylopectin

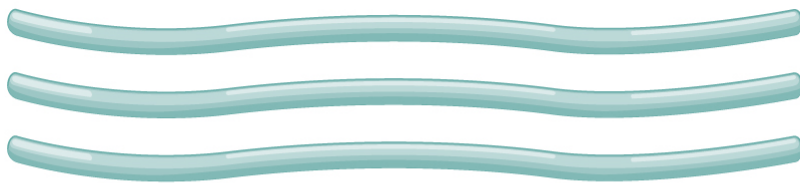


Glycogen is the storage form of glucose in humans

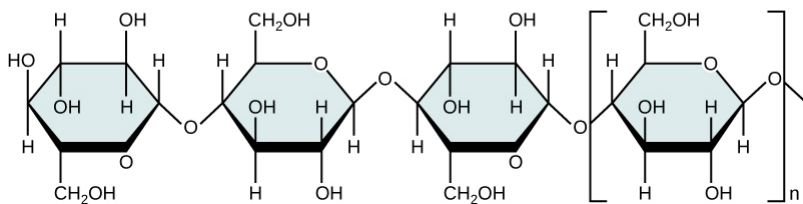
and other vertebrates and is comprised of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen breaks down to release glucose in a process scientists call glycogenolysis.

Cellulose is the most abundant natural biopolymer. Cellulose mostly comprises a plant's cell wall. This provides the cell structural support. Wood and paper are mostly cellulosic in nature. Glucose monomers comprise cellulose that β 1-4 glycosidic bonds link ([link]).

Cellulose fibers



Cellulose structure



As [link] shows, every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While human

digestive enzymes cannot break down the β 1-4 linkage, herbivores such as cows, koalas, and buffalos are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In some of these animals, certain species of bacteria and protists reside in the rumen (part of the herbivore's digestive system) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in ruminants' digestive systems. Cellulases can break down cellulose into glucose monomers that animals use as an energy source. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, the exoskeleton, which protects their internal body parts (as we see in the bee in [\[link\]](#)). This exoskeleton is made of the biological macromolecule **chitin**, which is a polysaccharide-containing nitrogen. It is made of repeating N-acetyl- β -d-glucosamine units, which are a modified sugar. Chitin is also a major component of fungal cell walls. Fungi are neither animals nor plants and form a kingdom of their own in the domain Eukarya.



Career Connections

Registered Dietitian

Obesity is a worldwide health concern, and many diseases such as diabetes and heart disease are becoming more prevalent because of obesity. This is one of the reasons why people increasingly seek out registered dietitians for advice. Registered dietitians help plan nutrition programs for individuals in various settings. They often work with patients in health care facilities, designing nutrition plans to treat and prevent diseases. For example, dietitians may teach a patient with diabetes how to manage blood sugar levels by

eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices.

To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam.

Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and physiology (biological functions) of food (proteins, carbohydrates, and fats).

Benefits of Carbohydrates

Are carbohydrates good for you? Some people believe that carbohydrates are bad and they should avoid them. Some diets completely forbid carbohydrate consumption, claiming that a low-carbohydrate diet helps people to lose weight faster. However, carbohydrates have been an important part of the human diet for thousands of years. Artifacts from ancient civilizations show the presence of wheat, rice, and corn in our ancestors' storage areas.

As part of a well balanced diet, we should supplement carbohydrates with proteins, vitamins, and fats. Calorie-wise, a gram of carbohydrate provides 4.3 Kcal. For comparison, fats provide 9 Kcal/g, a less desirable ratio. Carbohydrates contain soluble and insoluble elements. The insoluble part, fiber, is mostly cellulose. Fiber has many uses. It promotes regular bowel movement by adding bulk, and it regulates the blood glucose consumption rate. Fiber also helps to remove excess cholesterol from the body. Fiber binds to the cholesterol in the small intestine, then attaches to the cholesterol and prevents the cholesterol particles from entering the bloodstream. Cholesterol then exits the body via the feces. Fiber-rich diets also have a protective role in reducing the occurrence of colon cancer. In addition, a meal containing whole grains and vegetables gives a feeling of fullness. As an immediate source of energy, glucose breaks down during the cellular respiration process, which produces ATP, the cell's energy currency. Without consuming carbohydrates, we reduce the availability of “instant energy”. Eliminating carbohydrates from the diet may be necessary for some people, but such a step may not be healthy for everyone.

Link to Learning

For an additional perspective on carbohydrates, explore “Biomolecules: the Carbohydrates” through

this [interactive animation](#).

Section Summary

Carbohydrates are a group of macromolecules that are a vital energy source for the cell and provide structural support to plant cells, fungi, and all of the arthropods that include lobsters, crabs, shrimp, insects, and spiders. Scientists classify carbohydrates as monosaccharides, disaccharides, and polysaccharides depending on the number of monomers in the molecule. Monosaccharides are linked by glycosidic bonds that form as a result of dehydration reactions, forming disaccharides and polysaccharides with eliminating a water molecule for each bond formed. Glucose, galactose, and fructose are common monosaccharides; whereas, common disaccharides include lactose, maltose, and sucrose. Starch and glycogen, examples of polysaccharides, are the storage forms of glucose in plants and animals, respectively. The long polysaccharide chains may be branched or unbranched. Cellulose is an example of an unbranched polysaccharide; whereas, amylopectin, a constituent of starch, is a highly branched molecule. Glucose storage, in the form of polymers like starch or glycogen, makes it slightly less

accessible for metabolism; however, this prevents it from leaking out of the cell or creating a high osmotic pressure that could cause the cell to uptake excessive water.

Visual Connection Questions

[\[link\]](#) What kind of sugars are these, aldose or ketose?

[\[link\]](#) Glucose and galactose are aldoses. Fructose is a ketose.

Review Questions

An example of a monosaccharide is _____.

1. fructose
2. glucose
3. galactose
4. all of the above

Cellulose and starch are examples of:

1. monosaccharides
 2. disaccharides
 3. lipids
 4. polysaccharides
-

D

Plant cell walls contain which of the following in abundance?

1. starch
 2. cellulose
 3. glycogen
 4. lactose
-

B

Lactose is a disaccharide formed by the formation of a _____ bond between glucose and _____.

1. glycosidic; lactose
 2. glycosidic; galactose
 3. hydrogen; sucrose
 4. hydrogen; fructose
-

B

Which of the following is not an extracellular matrix role of carbohydrates?

1. protect an insect's internal organs from external trauma
2. prevent plant cells from lysing after the plant is watered
3. maintain the shape of a fungal spore
4. provide energy for muscle movement

D

Critical Thinking Questions

Describe the similarities and differences between glycogen and starch.

Glycogen and starch are polysaccharides. They are the storage form of glucose. Glycogen is stored in animals in the liver and in muscle cells, whereas starch is stored in the roots, seeds, and leaves of plants. Starch has two different forms, one unbranched (amylose) and

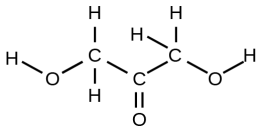
one branched (amylopectin), whereas glycogen is a single type of a highly branched molecule.

Why is it impossible for humans to digest food that contains cellulose?

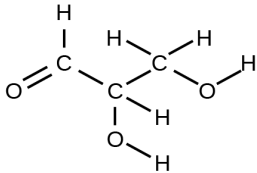
The β 1-4 glycosidic linkage in cellulose cannot be broken down by human digestive enzymes. Herbivores such as cows, koalas, and buffalos are able to digest grass that is rich in cellulose and use it as a food source because bacteria and protists in their digestive systems, especially in the rumen, secrete the enzyme cellulase. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal.

Draw the ketose and aldose forms of a monosaccharide with the chemical formula $C_3H_6O_3$. How is the structure of the monosaccharide changed from one form to the other in the human body?

The human body switches carbohydrates between their aldose and ketose forms using a family of enzymes called isomerases. The ketose triose is called dihydroxyacetone, and has the oxygen double-bonded to the center carbon:



The aldose is called glyceraldehyde, and can have the oxygen double-bonded to the first or third carbon of the molecule:



Glossary

carbohydrate

biological macromolecule in which the ratio of carbon to hydrogen and to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells and form arthropods' cellular exoskeleton

cellulose

polysaccharide that comprises the plants' cell wall; provides structural support to the cell

chitin

type of carbohydrate that forms the outer skeleton of all arthropods that include crustaceans and insects; it also forms fungi cell walls

disaccharide

two sugar monomers that a glycosidic bond links

glycogen

storage carbohydrate in animals

glycosidic bond

bond formed by a dehydration reaction between two monosaccharides with eliminating a water molecule

monosaccharide

single unit or monomer of carbohydrates

polysaccharide

long chain of monosaccharides; may be branched or unbranched

starch

storage carbohydrate in plants

Lipids

By the end of this section, you will be able to do the following:

- Describe the four major types of lipids
- Explain the role of fats in storing energy
- Differentiate between saturated and unsaturated fatty acids
- Describe phospholipids and their role in cells
- Define the basic structure of a steroid and some steroid functions
- Explain how cholesterol helps maintain the plasma membrane's fluid nature

Lipids include a diverse group of compounds that are largely nonpolar in nature. This is because they are hydrocarbons that include mostly nonpolar carbon–carbon or carbon–hydrogen bonds. Nonpolar molecules are hydrophobic (“water fearing”), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of fats. Lipids also provide insulation from the environment for plants and animals ([link]). For example, they help keep aquatic birds and mammals dry when forming a protective layer over fur or feathers because of their water-repellant hydrophobic nature. Lipids are also the building blocks of many hormones and are an important constituent of all cellular membranes. Lipids include fats, oils, waxes, phospholipids, and steroids. Hydrophobic lipids in aquatic mammals' fur, such as

this river otter, protect them from the elements.
(credit: Ken Bosma)



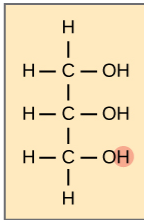
Joining three fatty acids to a glycerol backbone in a dehydration reaction forms triacylglycerol. Three water molecules release in the process. Stearic acid is a common saturated fatty acid. Oleic acid is a common unsaturated fatty acid. Saturated fatty acids have hydrocarbon chains connected by single bonds only. Unsaturated fatty acids have one or more double bonds. Each double bond may be in a *cis* or *trans* configuration. In the *cis* configuration, both hydrogens are on the same side of the hydrocarbon chain. In the *trans* configuration, the hydrogens are on opposite sides. A *cis* double bond causes a kink in the chain. Alpha-linolenic acid is an

example of an omega-3 fatty acid. It has three *cis* double bonds and, as a result, a curved shape. For clarity, the diagram does not show the carbons. Each singly bonded carbon has two hydrogens associated with it, which the diagram also does not show.

Fats and Oils

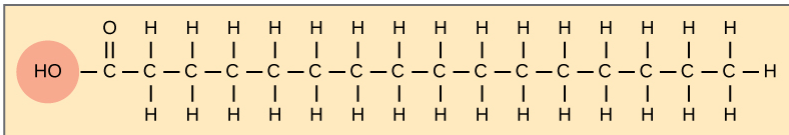
A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (alcohol) with three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name “fatty acid.” The number of carbons in the fatty acid may range from 4 to 36. The most common are those containing 12–18 carbons. In a fat molecule, the fatty acids attach to each of the glycerol molecule's three carbons with an ester bond through an oxygen atom ([\[link\]](#)).

Glycerol

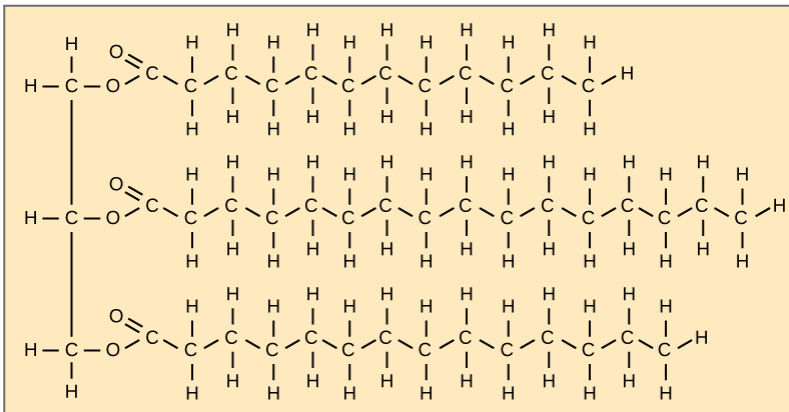


+

Fatty Acid



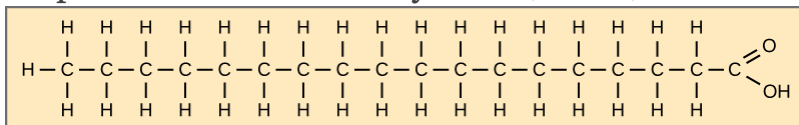
Triacylglycerol



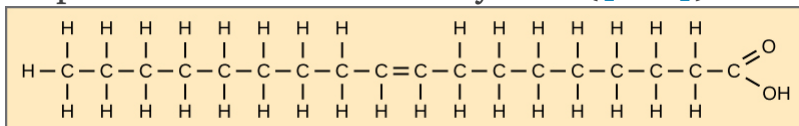
During this ester bond formation, three water molecules are released. The three fatty acids in the triacylglycerol may be similar or dissimilar. We also call fats **triacylglycerols** or **triglycerides** because of their chemical structure. Some fatty acids have common names that specify their origin. For example, palmitic acid, a **saturated fatty acid**, is

derived from the palm tree. Arachidic acid is derived from *Arachis hypogea*, the scientific name for groundnuts or peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is saturated. Saturated fatty acids are saturated with hydrogen. In other words, the number of hydrogen atoms attached to the carbon skeleton is maximized. Stearic acid is an example of a saturated fatty acid ([\[link\]](#)).



When the hydrocarbon chain contains a double bond, the fatty acid is **unsaturated**. Oleic acid is an example of an unsaturated fatty acid ([\[link\]](#)).



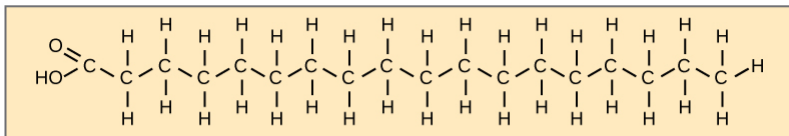
Most unsaturated fats are liquid at room temperature. We call these oils. If there is one double bond in the molecule, then it is a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is a polyunsaturated fat (e.g., canola oil).

When a fatty acid has no double bonds, it is a saturated fatty acid because it is not possible to add

more hydrogen to the chain's carbon atoms. A fat may contain similar or different fatty acids attached to glycerol. Long straight fatty acids with single bonds generally pack tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid (common in meat) and the fat with butyric acid (common in butter) are examples of saturated fats. Mammals store fats in specialized cells, or adipocytes, where fat globules occupy most of the cell's volume. Plants store fat or oil in many seeds and use them as a source of energy during seedling development. Unsaturated fats or oils are usually of plant origin and contain *cis* unsaturated fatty acids. *Cis* and *trans* indicate the configuration of the molecule around the double bond. If hydrogens are present in the same plane, it is a *cis* fat. If the hydrogen atoms are on two different planes, it is a **trans fat**. The *cis* double bond causes a bend or a “kink” that prevents the fatty acids from packing tightly, keeping them liquid at room temperature ([\[link\]](#)). Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to lower blood cholesterol levels; whereas, saturated fats contribute to plaque formation in the arteries.

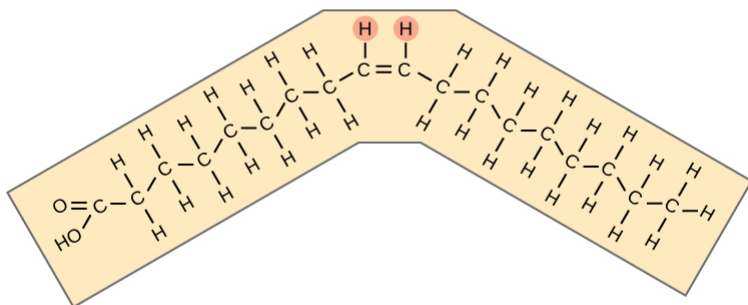
Saturated fatty acid

Stearic acid

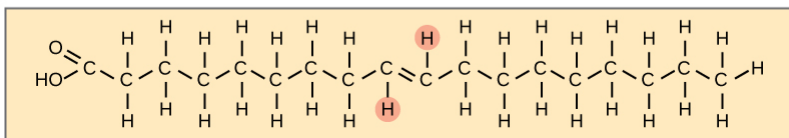


Unsaturated fatty acids

Cis oleic acid



Trans oleic acid



Trans Fats

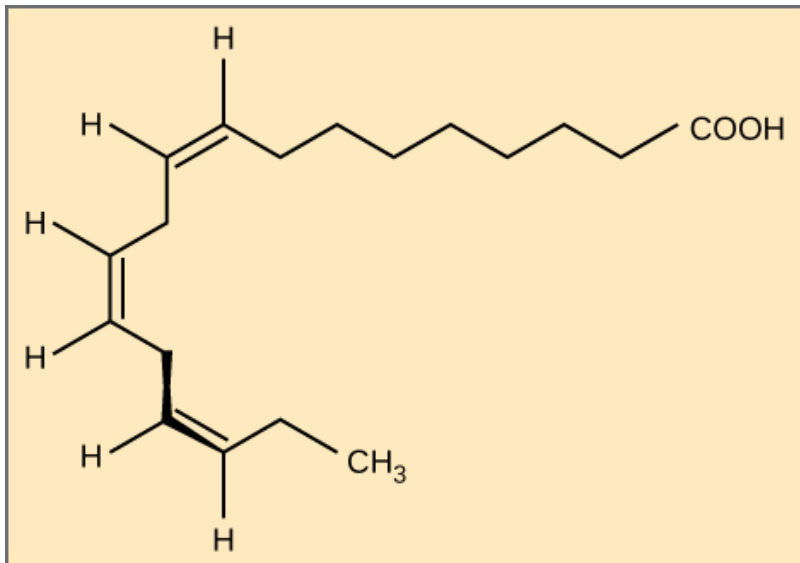
The food industry artificially hydrogenates oils to make them semi-solid and of a consistency desirable for many processed food products. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*- conformation in the hydrocarbon chain may convert to double bonds in the *trans*-conformation.

Margarine, some types of peanut butter, and

shortening are examples of artificially hydrogenated trans fats. Recent studies have shown that an increase in trans fats in the human diet may lead to higher levels of low-density lipoproteins (LDL), or “bad” cholesterol, which in turn may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently banned using trans fats, and food labels are required to display the trans fat content.

Omega Fatty Acids

Essential fatty acids are those that the human body requires but does not synthesize. Consequently, they have to be supplemented through ingestion via the diet. **Omega-3** fatty acids (like those in [\[link\]](#)) fall into this category and are one of only two known for humans (the other is omega-6 fatty acid). These are polyunsaturated fatty acids and are omega-3 because a double bond connects the third carbon from the hydrocarbon chain's end to its neighboring carbon.



The farthest carbon away from the carboxyl group is numbered as the omega (ω) carbon, and if the double bond is between the third and fourth carbon from that end, it is an omega-3 fatty acid.

Nutritionally important because the body does not make them, omega-3 fatty acids include alpha-linolenic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), all of which are polyunsaturated. Salmon, trout, and tuna are good sources of omega-3 fatty acids. Research indicates that omega-3 fatty acids reduce the risk of sudden death from heart attacks, lower triglycerides in the blood, decrease blood pressure, and prevent thrombosis by inhibiting blood clotting. They also reduce inflammation, and may help lower the risk of some cancers in animals.

Like carbohydrates, fats have received considerable

bad publicity. It is true that eating an excess of fried foods and other “fatty” foods leads to weight gain. However, fats do have important functions. Many vitamins are fat soluble, and fats serve as a long-term storage form of fatty acids: a source of energy. They also provide insulation for the body. Therefore, we should consume “healthy” fats in moderate amounts on a regular basis.

Lipids comprise waxy coverings on some leaves.
(credit: Roger Griffith)

Waxes

Wax covers some aquatic birds' feathers and some plants' leaf surfaces. Because of waxes' hydrophobic nature, they prevent water from sticking on the surface ([[link](#)]). Long fatty acid chains esterified to long-chain alcohols comprise waxes.

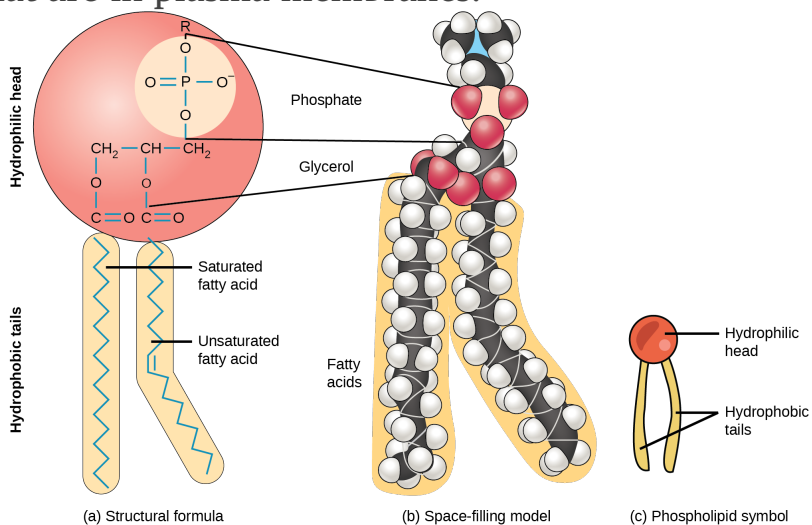


A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. Adding a charged or polar chemical group may modify the phosphate. The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

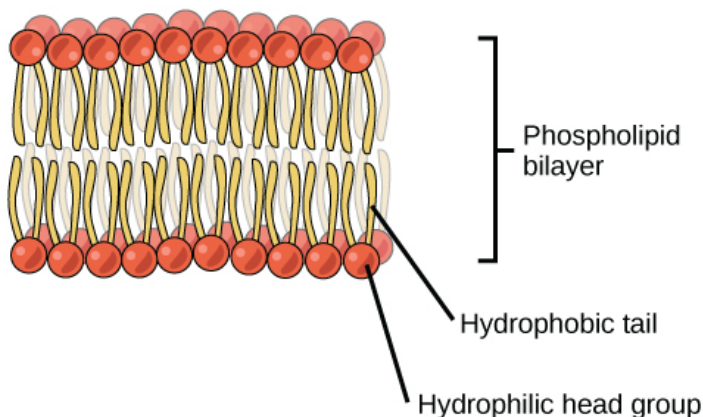
Phospholipids

Phospholipids are major plasma membrane constituents that comprise cells' outermost layer. Like fats, they are comprised of fatty acid chains attached to a glycerol or sphingosine backbone. However, instead of three fatty acids attached as in

triglycerides, there are two fatty acids forming diacylglycerol, and a modified phosphate group occupies the glycerol backbone's third carbon ([link]). A phosphate group alone attached to a diacylglycerol does not qualify as a phospholipid. It is phosphatidate (diacylglycerol 3-phosphate), the precursor of phospholipids. An alcohol modifies the phosphate group. Phosphatidylcholine and phosphatidylserine are two important phospholipids that are in plasma membranes.



A phospholipid is an amphipathic molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water; whereas, the phosphate-containing group is hydrophilic and interacts with water ([link]).



The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the structure's matrix, phospholipids' fatty acid tails face inside, away from water; whereas, the phosphate group faces the outside, aqueous side ([\[link\]](#)).

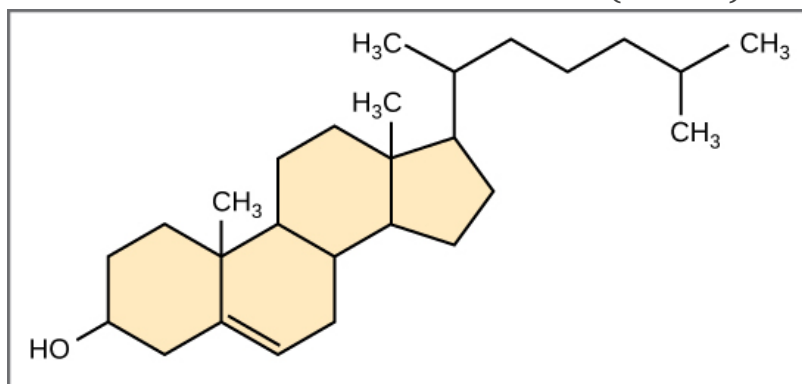
Phospholipids are responsible for the plasma membrane's dynamic nature. If a drop of phospholipids is placed in water, it spontaneously forms a structure that scientists call a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the structure's interior.

Four fused hydrocarbon rings comprise steroids such as cholesterol and cortisol.

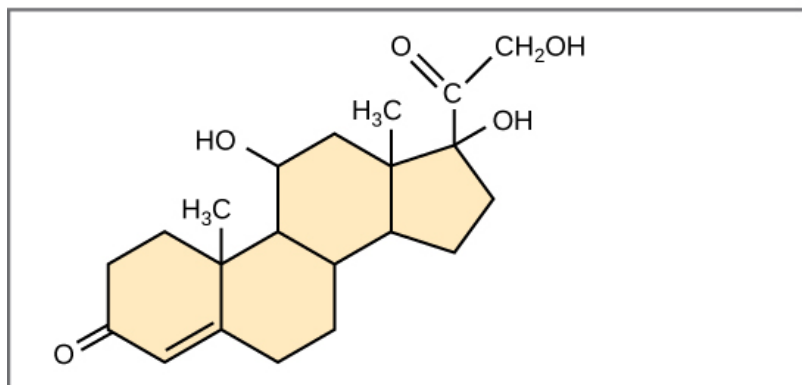
Steroids

Unlike the phospholipids and fats that we discussed earlier, **steroids** have a fused ring structure.

Although they do not resemble the other lipids, scientists group them with them because they are also hydrophobic and insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail ([link](#)). Many steroids also have the -OH functional group, which puts them in the alcohol classification (sterols).



Cholesterol



Cortisol

Cholesterol is the most common steroid. The liver synthesizes cholesterol and is the precursor to many steroid hormones such as testosterone and estradiol,

which gonads and endocrine glands secrete. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help emulsifying fats and their subsequent absorption by cells.

Although lay people often speak negatively about cholesterol, it is necessary for the body's proper functioning. Sterols (cholesterol in animal cells, phytosterol in plants) are components of the plasma membrane of cells and are found within the phospholipid bilayer.

Link to Learning

For an additional perspective on lipids, explore the interactive animation “[Biomolecules: The Lipids](#)”.

Section Summary

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats are a stored form of energy and are also known as triacylglycerols or triglycerides. Fats are comprised of fatty acids and either glycerol or sphingosine. Fatty acids may be unsaturated or saturated, depending on the presence or absence of

double bonds in the hydrocarbon chain. If only single bonds are present, they are saturated fatty acids. Unsaturated fatty acids may have one or more double bonds in the hydrocarbon chain.

Phospholipids comprise the membrane's matrix. They have a glycerol or sphingosine backbone to which two fatty acid chains and a phosphate-containing group are attached. Steroids are another class of lipids. Their basic structure has four fused carbon rings. Cholesterol is a type of steroid and is an important constituent of the plasma membrane, where it helps to maintain the membrane's fluid nature. It is also the precursor of steroid hormones such as testosterone.

Review Questions

Saturated fats have all of the following characteristics except:

1. they are solid at room temperature
2. they have single bonds within the carbon chain
3. they are usually obtained from animal sources
4. they tend to dissolve in water easily

Phospholipids are important components of _____.

1. the plasma membrane of cells
2. the ring structure of steroids
3. the waxy covering on leaves
4. the double bond in hydrocarbon chains

A

Cholesterol is an integral part of plasma membranes. Based on its structure, where is it found in the membrane?

1. on the extracellular surface
2. embedded with the phospholipid heads
3. within the tail bilayer
4. attached to the intracellular surface

C

Critical Thinking Questions

Explain at least three functions that lipids serve

in plants and/or animals.

Fat serves as a valuable way for animals to store energy. It can also provide insulation. Waxes can protect plant leaves and mammalian fur from getting wet. Phospholipids and steroids are important components of animal cell membranes, as well as plant, fungal, and bacterial membranes.

Why have trans fats been banned from some restaurants? How are they created?

Trans fats are created artificially when hydrogen gas is bubbled through oils to solidify them. The double bonds of the *cis* conformation in the hydrocarbon chain may be converted to double bonds in the *trans* configuration. Some restaurants are banning trans fats because they cause higher levels of LDL, or “bad” cholesterol.

Why are fatty acids better than glycogen for storing large amounts of chemical energy?

Fats have a higher energy density than carbohydrates (averaging 9kcal/gram versus 4.3kcal/gram respectively). Thus, on a per gram

basis, more energy can be stored in fats than can be stored in carbohydrates. Additionally, fats are packaged into spherical globules to minimize interactions with the water-based plasma membrane, while glycogen is a large branched carbohydrate that cannot be compacted for storage.

Part of cortisol's role in the body involves passing through the plasma membrane to initiate signaling inside a cell. Describe how the structures of cortisol and the plasma membrane allow this to occur.

Cortisol is a small, generally hydrophobic molecule, while the phospholipids that create plasma membranes have a hydrophilic head and hydrophobic tails. Since cortisol is hydrophobic, it can interact with the sequestered tails of the phospholipids in the center of the plasma membrane. This, along with its small size, allows cortisol to move through the plasma membrane to the inside of the cell.

Glossary

lipid

macromolecule that is nonpolar and insoluble

in water

omega fat

type of polyunsaturated fat that the body requires; numbering the carbon omega starts from the methyl end or the end that is farthest from the carboxylic end

phospholipid

membranes' major constituent; comprised of two fatty acids and a phosphate-containing group attached to a glycerol backbone

saturated fatty acid

long-chain hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

steroid

type of lipid comprised of four fused hydrocarbon rings forming a planar structure

trans fat

fat formed artificially by hydrogenating oils, leading to a different arrangement of double bond(s) than those in naturally occurring lipids

triacylglycerol (also, triglyceride)

fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid

long-chain hydrocarbon that has one or more double bonds in the hydrocarbon chain

wax

lipid comprised of a long-chain fatty acid that is esterified to a long-chain alcohol; serves as a protective coating on some feathers, aquatic mammal fur, and leaves

Proteins

By the end of this section, you will be able to do the following:

- Describe the functions proteins perform in the cell and in tissues
- Discuss the relationship between amino acids and proteins
- Explain the four levels of protein organization
- Describe the ways in which protein shape and function are linked

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective. They may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, amino acid polymers arranged in a linear sequence.

Types and Functions of Proteins

Enzymes, which living cells produce, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each

enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. The enzyme may help in breakdown, rearrangement, or synthesis reactions. We call enzymes that break down their substrates catabolic enzymes. Those that build more complex molecules from their substrates are anabolic enzymes, and enzymes that affect the rate of reaction are catalytic enzymes. Note that all enzymes increase the reaction rate and, therefore, are organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

Hormones are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps regulate the blood glucose level. [\[link\]](#) lists the primary types and functions of proteins.

Protein Types and Functions		
Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in food by catabolizing

			nutrients into monomeric units
Transport	Hemoglobin, albumin		Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin		Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine		Coordinate different body systems' activity
Defense	Immunoglobulins		Protect the body from foreign pathogens
Contractile	Actin, myosin		Effect muscle contraction
Storage	Legume storage proteins, egg white (albumin)		Provide nourishment in early embryo development and the seedling

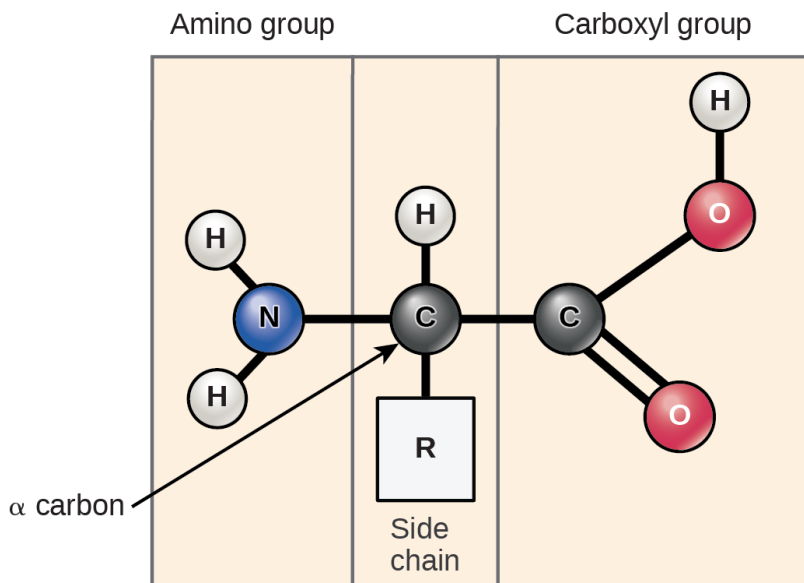
Proteins have different shapes and molecular weights. Some proteins are globular in shape; whereas, others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, located in our skin, is a fibrous protein. Protein shape is critical to its function, and many different

types of chemical bonds maintain this shape. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the protein's shape, leading to loss of function, or **denaturation**. Different arrangements of the same 20 types of amino acids comprise all proteins. Two rare new amino acids were discovered recently (selenocystein and pirrolysine), and additional new discoveries may be added to the list.

Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached. Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the incoming amino acid's amino group. In the process, it releases a water molecule.

Amino Acids

Amino acids are the monomers that comprise proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, or the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group ([\[link\]](#)).



Scientists use the name "amino acid" because these acids contain both amino group and carboxyl-acid-group in their basic structure. As we mentioned, there are 20 common amino acids present in proteins. Nine of these are essential amino acids in humans because the human body cannot produce them and we obtain them from our diet. For each amino acid, the R group (or side chain) is different ([\[link\]](#)).

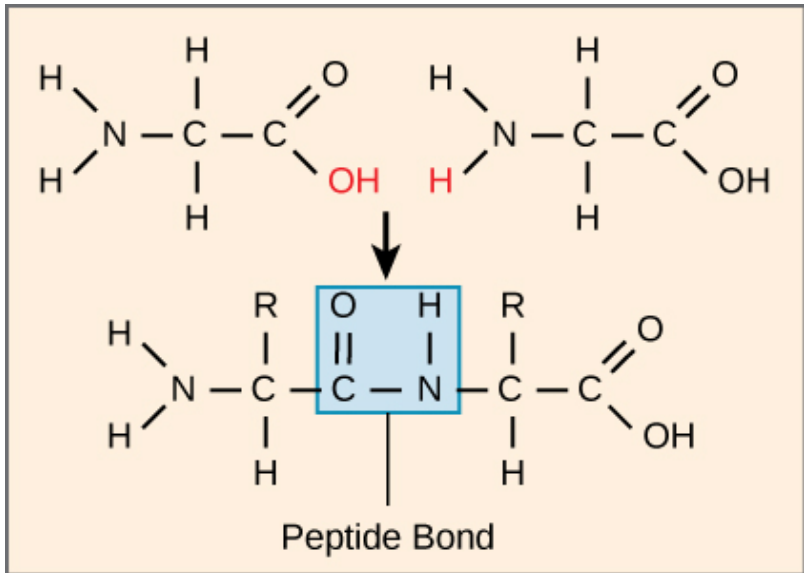
Visual Connection

There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the amino acid's standard structure since its amino group is not separate from the side chain ([\[link\]](#)).

A single upper case letter or a three-letter abbreviation represents amino acids. For example, the letter V or the three-letter symbol val represent valine. Just as some fatty acids are essential to a diet, some amino acids also are necessary. These essential amino acids in humans include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary to build proteins in the body, but not those that the body produces. Which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. A covalent bond, or **peptide bond**, attaches to each amino acid, which a dehydration reaction forms. One amino acid's carboxyl group and the incoming amino acid's amino group combine, releasing a water molecule. The resulting bond is the peptide bond ([\[link\]](#)).



The products that such linkages form are peptides. As more amino acids join to this growing chain, the resulting chain is a polypeptide. Each polypeptide has a free amino group at one end. This end is the N terminal, or the amino terminal, and the other end has a free carboxyl group, also the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require adding other chemical groups. Only after these modifications is

the protein completely functional.

Link to Learning

Click through the steps of protein synthesis in this [interactive tutorial](#).

Evolution Connection

The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally located in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the heme's central ion alternately reduces and oxidizes during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species. In other words, we can assess evolutionary kinship by measuring the similarities or differences among various species' DNA or protein sequences.

Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that scientists have sequenced to date, 37 of these amino acids

appear in the same position in all cytochrome c samples. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, scientists did not find a sequence difference. When researchers compared human and rhesus monkey sequences, the single difference was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

Bovine serum insulin is a protein hormone comprised of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, three-letter abbreviations that represent the amino acids' names in the order they are present indicate primary structure. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but we have drawn them different sizes for clarity. The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, a valine replaces glutamate. In this blood smear, visualized at 535x magnification using

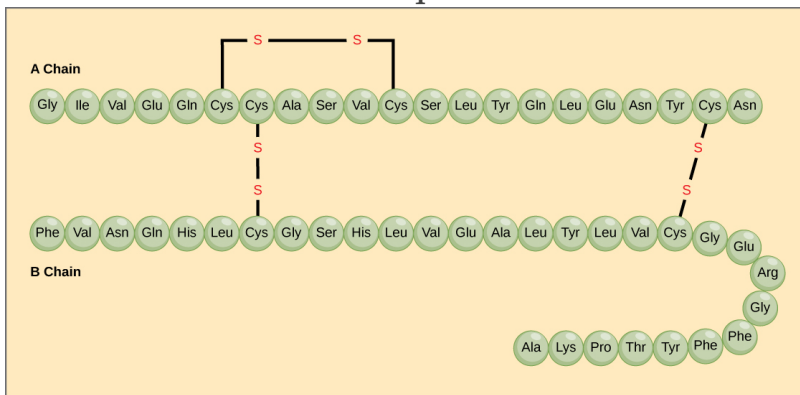
bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell) The α -helix and β -pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an α -helix, while others have a propensity to form a β -pleated sheet. A variety of chemical interactions determine the proteins' tertiary structure. These include hydrophobic interactions, ionic bonding, hydrogen bonding, and disulfide linkages. Observe the four levels of protein structure in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Protein Structure

As we discussed earlier, a protein's shape is critical to its function. For example, an enzyme can bind to a specific substrate at an active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

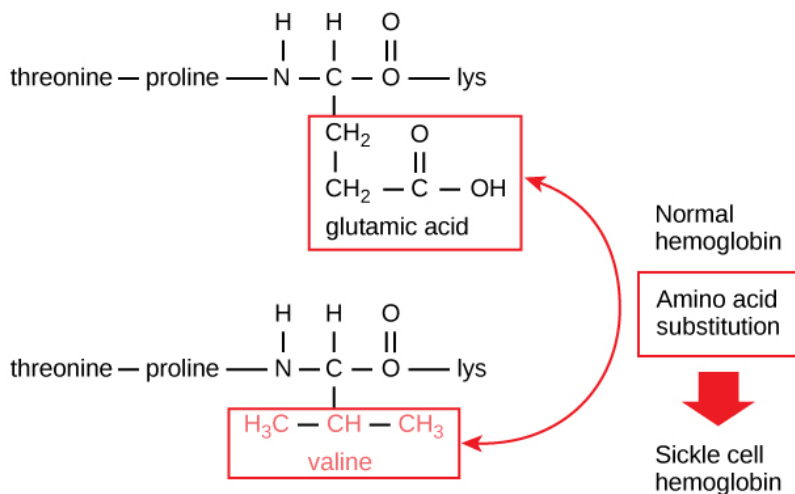
Primary Structure

Amino acids' unique sequence in a polypeptide chain is its **primary structure**. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine; whereas, the C terminal amino acid is asparagine ([\[link\]](#)). The amino acid sequences in the A and B chains are unique to insulin.

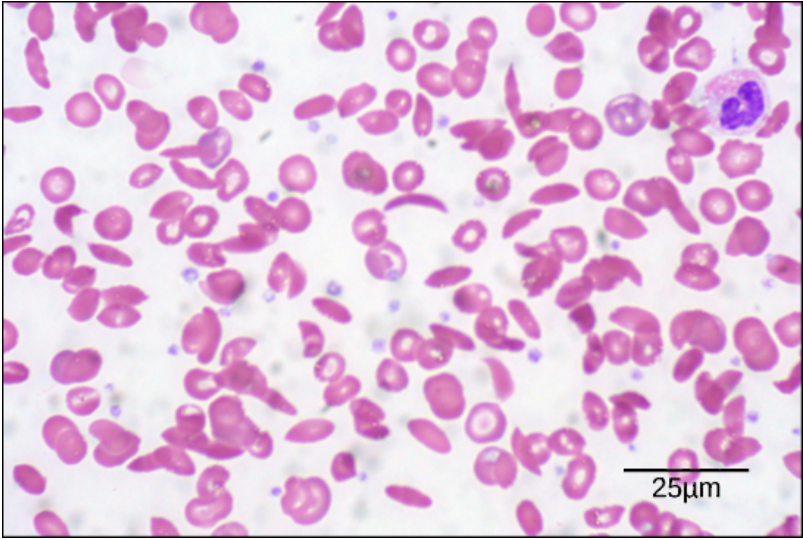


The gene encoding the protein ultimately determines the unique sequence for every protein. A change in nucleotide sequence of the gene's coding region may lead to adding a different amino acid to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain (a small portion of which we show in [\[link\]](#)) has a single amino acid substitution, causing a change in protein structure and function. Specifically, valine in the β chain substitutes the amino acid glutamic. What is most remarkable to consider is that a hemoglobin molecule is comprised of two alpha and two beta chains that each consist of about 150 amino acids.

The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that three nucleotides each encode those 600 amino acids, and a single base change (point mutation), 1 in 1800 bases causes the mutation.



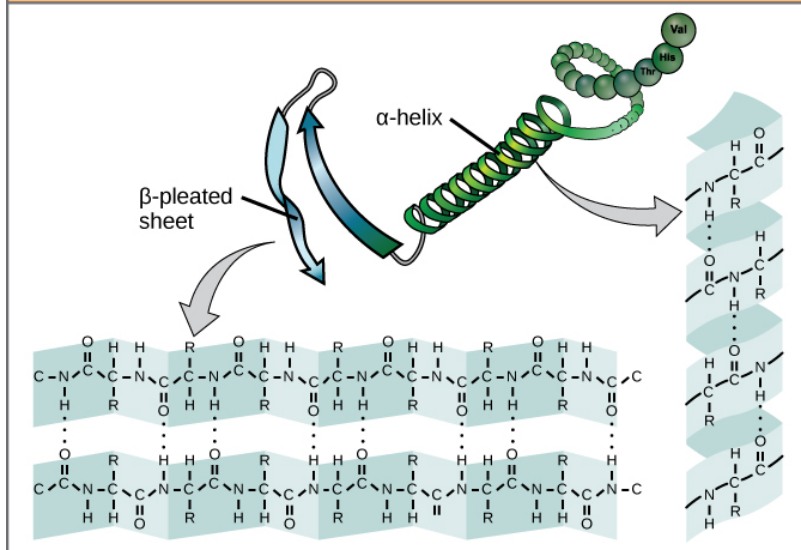
Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and causes them to assume a crescent or “sickle” shape, which clogs blood vessels ([\[link\]](#)). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.



Secondary Structure

The local folding of the polypeptide in some regions gives rise to the **secondary structure** of the protein. The most common are the **α -helix** and **β -pleated sheet** structures ([\[link\]](#)). Both structures are held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.

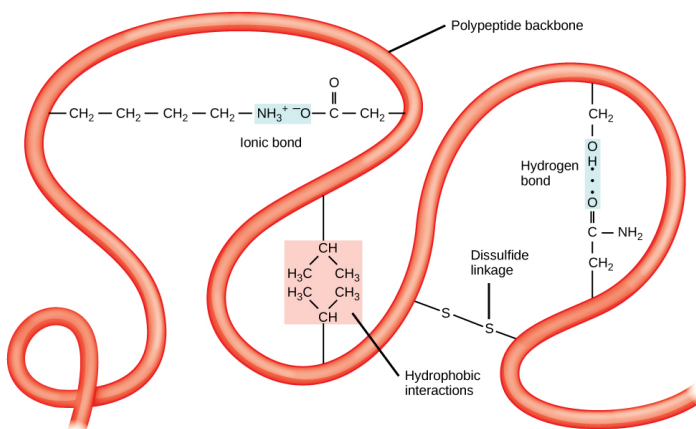
Secondary Protein Structure



Every helical turn in an alpha helix has 3.6 amino acid residues. The polypeptide's R groups (the variant groups) protrude out from the α -helix chain. In the β -pleated sheet, hydrogen bonding between atoms on the polypeptide chain's backbone form the "pleats". The R groups are attached to the carbons and extend above and below the pleat's folds. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially negative oxygen atom in the peptide backbone's carbonyl group. The α -helix and β -pleated sheet structures are in most globular and fibrous proteins and they play an important structural role.

Tertiary Structure

The polypeptide's unique three-dimensional structure is its **tertiary structure** ([\[link\]](#)). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups create the protein's complex three-dimensional tertiary structure. The nature of the R groups in the amino acids involved can counteract forming the hydrogen bonds we described for standard secondary structures. For example, R groups with like charges repel each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the nonpolar amino acids' hydrophobic R groups lie in the protein's interior; whereas, the hydrophilic R groups lie on the outside. Scientists also call the former interaction types hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond that forms during protein folding.

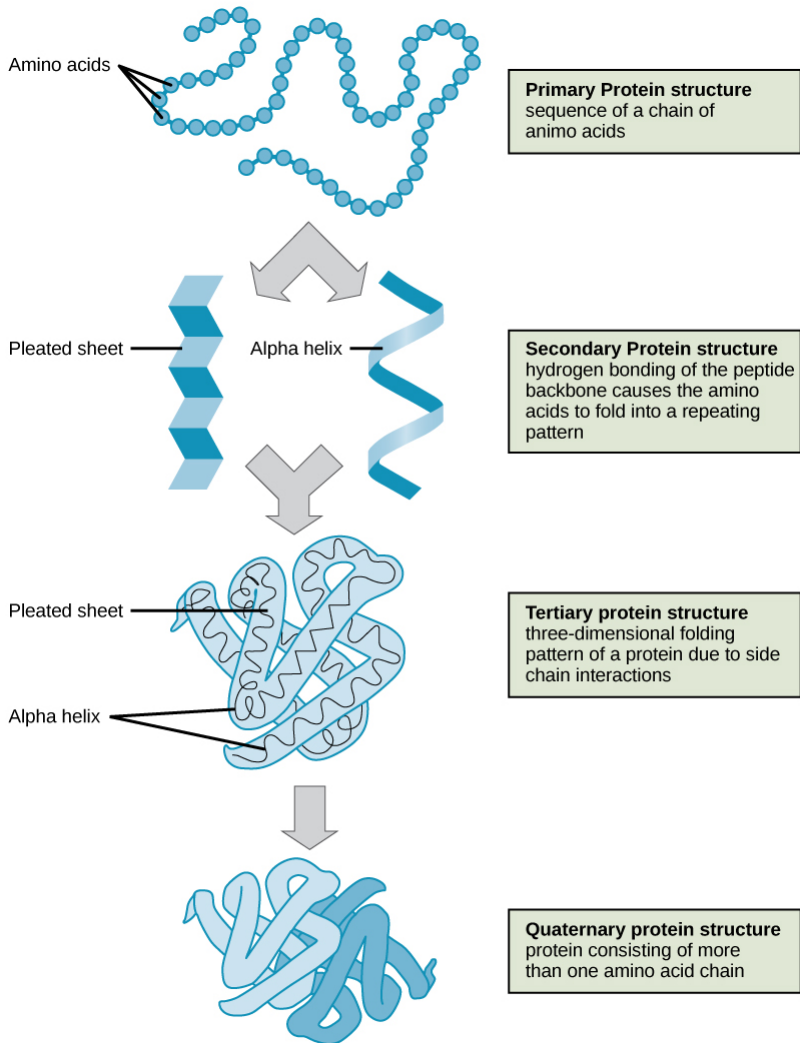


All of these interactions, weak and strong, determine the protein's final three-dimensional shape. When a protein loses its three-dimensional shape, it may no longer be functional.

Quaternary Structure

In nature, some proteins form from several polypeptides, or subunits, and the interaction of these subunits forms the **quaternary structure**. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen and disulfide bonds that cause it to mostly clump into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after forming the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β -pleated sheet structure that is the result of hydrogen bonding between different chains.

[\[link\]](#) illustrates the four levels of protein structure (primary, secondary, tertiary, and quaternary).



Denaturation and Protein Folding

Each protein has its own unique sequence and shape that chemical interactions hold together. If the protein is subject to changes in temperature, pH, or

exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what scientists call denaturation. Denaturation is often reversible because the polypeptide's primary structure is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is frying an egg. The albumin protein in the liquid egg white denatures when placed in a hot pan. Not all proteins denature at high temperatures. For instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the stomach's digestive enzymes retain their activity under these conditions.

Protein folding is critical to its function. Scientists originally thought that the proteins themselves were responsible for the folding process. Only recently researchers discovered that often they receive assistance in the folding process from protein helpers, or **chaperones** (or chaperonins) that associate with the target protein during the folding process. They act by preventing polypeptide aggregation that comprise the complete protein structure, and they disassociate from the protein once the target protein is folded.

Link to Learning

For an additional perspective on proteins, view [this animation](#) called “Biomolecules: The Proteins.”

Section Summary

Proteins are a class of macromolecules that perform a diverse range of functions for the cell. They help in metabolism by acting as enzymes, carriers, or hormones, and provide structural support. The building blocks of proteins (monomers) are amino acids. Each amino acid has a central carbon that bonds to an amino group, a carboxyl group, a hydrogen atom, and an R group or side chain. There are 20 commonly occurring amino acids, each of which differs in the R group. A peptide bond links each amino acid to its neighbors. A long amino acid chain is a polypeptide.

Proteins are organized at four levels: primary, secondary, tertiary, and (optional) quaternary. The primary structure is the amino acids' unique sequence. The polypeptide's local folding to form structures such as the α -helix and β -pleated sheet constitutes the secondary structure. The overall three-dimensional structure is the tertiary structure. When two or more polypeptides combine to form

the complete protein structure, the configuration is the protein's quaternary structure. Protein shape and function are intricately linked. Any change in shape caused by changes in temperature or pH may lead to protein denaturation and a loss in function.

Visual Connection Questions

[\[link\]](#) Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

[\[link\]](#) Polar and charged amino acid residues (the remainder after peptide bond formation) are more likely to be found on the surface of soluble proteins where they can interact with water, and nonpolar (e.g., amino acid side chains) are more likely to be found in the interior where they are sequestered from water. In membrane proteins, nonpolar and hydrophobic amino acid side chains associate with the hydrophobic tails of phospholipids, while polar and charged amino acid side chains interact with the polar head groups or with the

aqueous solution. However, there are exceptions. Sometimes, positively and negatively charged amino acid side chains interact with one another in the interior of a protein, and polar or charged amino acid side chains that interact with a ligand can be found in the ligand binding pocket.

Review Questions

The monomers that make up proteins are called _____.

1. nucleotides
2. disaccharides
3. amino acids
4. chaperones

C

The α -helix and the β -pleated sheet are part of which protein structure?

1. primary
2. secondary
3. tertiary

4. quaternary

B

Mad cow disease is an infectious disease where one misfolded protein causes all other copies of the protein to begin misfolding. This is an example of a disease impacting ___ structure.

1. primary
 2. secondary
 3. tertiary
 4. quaternary
-

C

Critical Thinking Questions

Explain what happens if even one amino acid is substituted for another in a polypeptide chain. Provide a specific example.

A change in gene sequence can lead to a different amino acid being added to a polypeptide chain instead of the normal one.

This causes a change in protein structure and function. For example, in sickle cell anemia, the hemoglobin β chain has a single amino acid substitution—the amino acid glutamic acid in position six is substituted by valine. Because of this change, hemoglobin molecules form aggregates, and the disc-shaped red blood cells assume a crescent shape, which results in serious health problems.

Describe the differences in the four protein structures.

The sequence and number of amino acids in a polypeptide chain is its primary structure. The local folding of the polypeptide in some regions is the secondary structure of the protein. The three-dimensional structure of a polypeptide is known as its tertiary structure, created in part by chemical interactions such as hydrogen bonds between polar side chains, van der Waals interactions, disulfide linkages, and hydrophobic interactions. Some proteins are formed from multiple polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure.

Aquaporins are proteins embedded in the plasma membrane that allow water molecules

to move between the extracellular matrix and the intracellular space. Based on its function and location, describe the key features of the protein's shape and the chemical characteristics of its amino acids.

The protein must form a channel in the plasma membrane that allows water into the cell since water cannot cross the plasma membrane by itself. Since aquaporins are embedded in the plasma membrane and connect with both the intracellular and extracellular spaces, it must be amphipathic like the plasma membrane. The top and bottom of the protein must contain charged or polar amino acids (hydrophilic) to interact with the aqueous environments. The exterior transmembrane region must contain non-polar amino acids (hydrophobic) that can interact with the phospholipid tails. However, the inside of this channel must contain hydrophilic amino acids since they will interact with the traveling water molecules.

Glossary

alpha-helix structure (α -helix)

type of secondary protein structure formed by folding the polypeptide into a helix shape with hydrogen bonds stabilizing the structure

amino acid

a protein's monomer; has a central carbon or alpha carbon to which an amino group, a carboxyl group, a hydrogen, and an R group or side chain is attached; the R group is different for all 20 common amino acids

beta-pleated sheet (β -pleated)

secondary structure in proteins in which hydrogen bonding forms "pleats" between atoms on the polypeptide chain's backbone

chaperone

(also, chaperonin) protein that helps nascent protein in the folding process

denaturation

loss of shape in a protein as a result of changes in temperature, pH, or chemical exposure

enzyme

catalyst in a biochemical reaction that is usually a complex or conjugated protein

hormone

chemical signaling molecule, usually protein or steroid, secreted by endocrine cells that act to control or regulate specific physiological processes

peptide bond

bond formed between two amino acids by a dehydration reaction

polypeptide

long chain of amino acids that peptide bonds link

primary structure

linear sequence of amino acids in a protein

protein

biological macromolecule comprised of one or more amino acid chains

quaternary structure

association of discrete polypeptide subunits in a protein

secondary structure

regular structure that proteins form by intramolecular hydrogen bonding between the oxygen atom of one amino acid residue and the hydrogen attached to the nitrogen atom of another amino acid residue

tertiary structure

a protein's three-dimensional conformation, including interactions between secondary structural elements; formed from interactions between amino acid side chains

Nucleic Acids

By the end of this section, you will be able to do the following:

- Describe nucleic acids' structure and define the two types of nucleic acids
- Explain DNA's structure and role
- Explain RNA's structure and roles

Nucleic acids are the most important macromolecules for the continuity of life. They carry the cell's genetic blueprint and carry instructions for its functioning.

Three components comprise a nucleotide: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1' through 5' (the prime distinguishes these residues from those in the base, which are numbered without using a prime notation). The base is attached to the ribose's 1' position, and the phosphate is attached to the 5' position. When a polynucleotide forms, the incoming nucleotide's 5' phosphate attaches to the 3' hydroxyl group at the end of the growing chain. Two types of pentose are in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2' position. We can divide bases into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

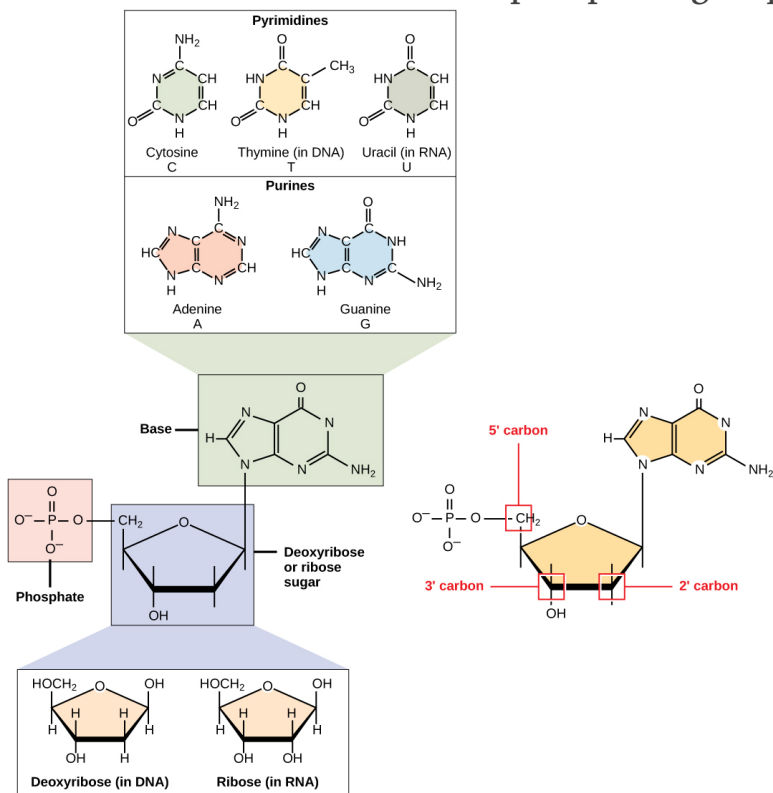
DNA and RNA

The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The cell's entire genetic content is its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products. Other genes code for RNA products. DNA controls all of the cellular activities by turning the genes “on” or “off.”

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the **messenger RNA (mRNA)**. Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are comprised of monomers that scientists call **nucleotides**. The nucleotides combine with each other to form a **polynucleotide**, DNA or RNA. Three components comprise each nucleotide: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group ([\[link\]](#)). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.



The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus

decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Scientists classify adenine and guanine as **purines**. The purine's primary structure is two carbon-nitrogen rings. Scientists classify cytosine, thymine, and uracil as **pyrimidines** which have a single carbon-nitrogen ring as their primary structure ([\[link\]](#)). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, we know the nitrogenous bases by their symbols A, T, G, C, and U. DNA contains A, T, G, and C; whereas, RNA contains A, U, G, and C.

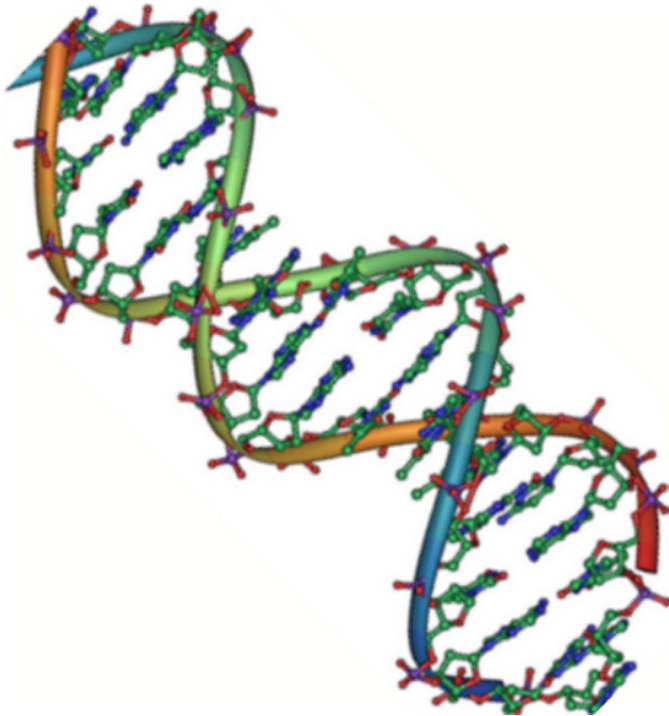
The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose ([\[link\]](#)). The difference between the sugars is the presence of the hydroxyl group on the ribose's second carbon and hydrogen on the deoxyribose's second carbon. The carbon atoms of the sugar molecule are numbered as 1', 2', 3', 4', and 5' (1' is read as "one prime"). The phosphate residue attaches to the hydroxyl group of the 5' carbon of one sugar and the hydroxyl group of the 3' carbon of the sugar of the next nucleotide, which forms a 5'–3' **phosphodiester** linkage. A simple dehydration reaction like the other linkages connecting monomers in macromolecules does not

form the phosphodiester linkage. Its formation involves removing two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand.
(credit: Jerome Walker/Dennis Myts)

DNA Double-Helix Structure

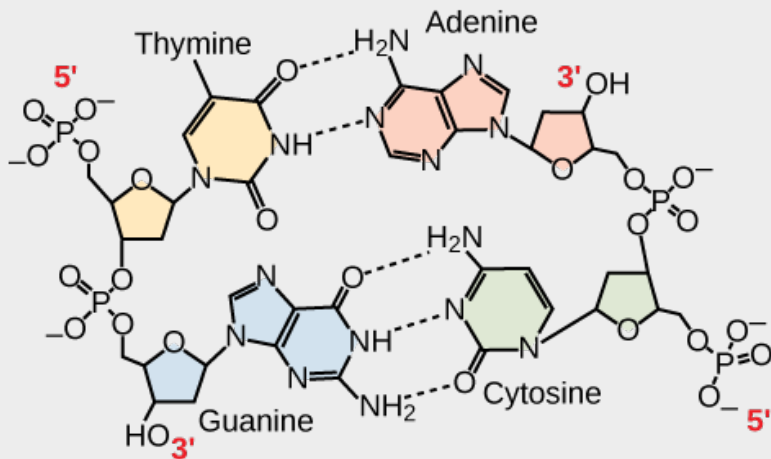
DNA has a double-helix structure ([\[link\]](#)). The sugar and phosphate lie on the outside of the helix, forming the DNA's backbone. The nitrogenous bases are stacked in the interior, like a pair of staircase steps. Hydrogen bonds bind the pairs to each other. Every base pair in the double helix is separated from the next base pair by 0.34 nm. The helix's two strands run in opposite directions, meaning that the 5' carbon end of one strand will face the 3' carbon end of its matching strand. (Scientists call this an antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)



Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as [\[link\]](#) shows. This is the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand copies itself, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.

Visual Connection

In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5' to 3' and the other 3' to 5'. The phosphate backbone is located on the outside, and the bases are in the middle. Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.



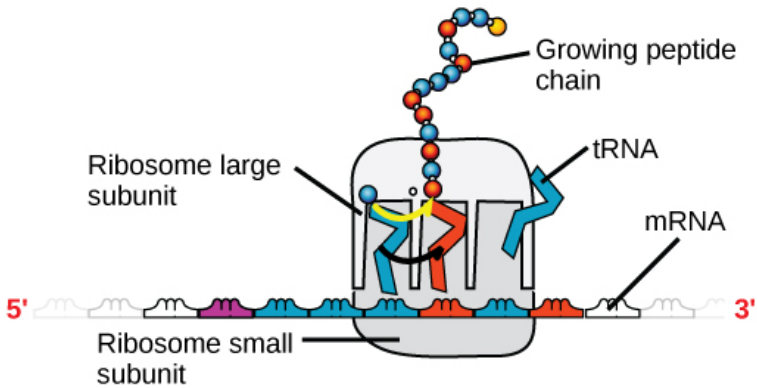
A mutation occurs, and adenine replaces cytosine. What impact do you think this will have on the DNA structure?

A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is comprised of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires synthesizing a certain protein, the gene for this product turns “on” and the messenger RNA synthesizes in the nucleus. The RNA base sequence is complementary to the DNA's coding sequence from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery ([\[link\]](#)).



The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. **Ribosomal RNA (rRNA)** is a major constituent of ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the Ribosomes. The ribosome's rRNA also has an enzymatic activity (peptidyl transferase) and catalyzes peptide bond formation between two aligned amino acids. **Transfer RNA (tRNA)** is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the protein synthesis site. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to insert itself in the polypeptide chain. MicroRNAs are the smallest RNA molecules and their role involves regulating gene expression by interfering with the expression of certain mRNA messages. [\[link\]](#) summarizes DNA and RNA features.

DNA and RNA		
Features		
	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process scientists call **transcription**, and RNA dictates the protein's structure in a process scientists call **translation**. This is the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

Link to Learning

To learn more about DNA, explore the [Howard Hughes Medical Institute BioInteractive animations](#) on the topic of DNA.

Section Summary

Nucleic acids are molecules comprised of nucleotides that direct cellular activities such as cell division and protein synthesis. Pentose sugar, a nitrogenous base, and a phosphate group comprise each nucleotide. There are two types of nucleic acids: DNA and RNA. DNA carries the cell's genetic blueprint and passes it on from parents to offspring (in the form of chromosomes). It has a double-helical structure with the two strands running in opposite directions, connected by hydrogen bonds, and complementary to each other. RNA is a single-stranded polymer composed of linked nucleotides made up of a pentose sugar (ribose), a nitrogenous base, and a phosphate group. RNA is involved in protein synthesis and its regulation. Messenger RNA (mRNA) copies from the DNA, exports itself from the nucleus to the cytoplasm, and contains information for constructing proteins. Ribosomal RNA (rRNA) is a part of the ribosomes at the site of

protein synthesis; whereas, transfer RNA (tRNA) carries the amino acid to the site of protein synthesis. The microRNA regulates using mRNA for protein synthesis.

Visual Connection Questions

[\[link\]](#) A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

[\[link\]](#) Adenine is larger than cytosine and will not be able to base pair properly with the guanine on the opposing strand. This will cause the DNA to bulge. DNA repair enzymes may recognize the bulge and replace the incorrect nucleotide.

Review Questions

A nucleotide of DNA may contain _____.

1. ribose, uracil, and a phosphate group
2. deoxyribose, uracil, and a phosphate group

3. deoxyribose, thymine, and a phosphate group
 4. ribose, thymine, and a phosphate group
-

C

The building blocks of nucleic acids are _____.

1. sugars
 2. nitrogenous bases
 3. peptides
 4. nucleotides
-

D

How does the double helix structure of DNA support its role in encoding the genome?

1. The sugar-phosphate backbone provides a template for DNA replication.
 2. tRNA pairing with the template strand creates proteins encoded by the genome.
 3. Complementary base pairing creates a very stable structure.
 4. Complementary base pairing allows for easy editing of both strands of DNA.
-

Critical Thinking Questions

What are the structural differences between RNA and DNA?

DNA has a double-helix structure. The sugar and the phosphate are on the outside of the helix and the nitrogenous bases are in the interior. The monomers of DNA are nucleotides containing deoxyribose, one of the four nitrogenous bases (A, T, G and C), and a phosphate group. RNA is usually single-stranded and is made of ribonucleotides that are linked by phosphodiester linkages. A ribonucleotide contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

What are the four types of RNA and how do they function?

The four types of RNA are messenger RNA, ribosomal RNA, transfer RNA, and microRNA.

Messenger RNA carries the information from the DNA that controls all cellular activities. The mRNA binds to the ribosomes that are constructed of proteins and rRNA, and tRNA transfers the correct amino acid to the site of protein synthesis. microRNA regulates the availability of mRNA for translation.

Glossary

deoxyribonucleic acid (DNA)

double-helical molecule that carries the cell's hereditary information

messenger RNA (mRNA)

RNA that carries information from DNA to ribosomes during protein synthesis

nucleic acid

biological macromolecule that carries the cell's genetic blueprint and carries instructions for the cell's functioning

nucleotide

monomer of nucleic acids; contains a pentose sugar, one or more phosphate groups, and a nitrogenous base

phosphodiester

linkage covalent chemical bond that holds together the polynucleotide chains with a

phosphate group linking neighboring nucleotides' two pentose sugars

polynucleotide

long chain of nucleotides

purine

type of nitrogenous base in DNA and RNA; adenine and guanine are purines

pyrimidine

type of nitrogenous base in DNA and RNA; cytosine, thymine, and uracil are pyrimidines

ribonucleic acid (RNA)

single-stranded, often internally base paired, molecule that is involved in protein synthesis

ribosomal RNA (rRNA)

RNA that ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes forming the peptide linkage

transcription

process through which messenger RNA forms on a template of DNA

transfer RNA (tRNA)

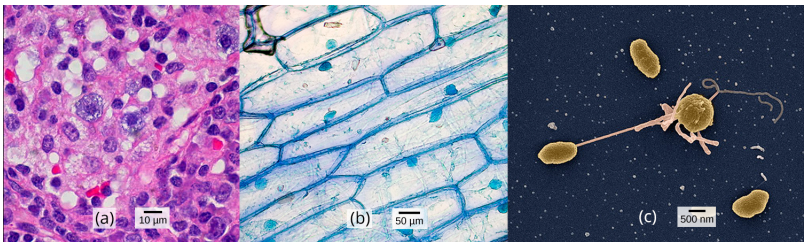
RNA that carries activated amino acids to the site of protein synthesis on the ribosome

translation

process through which RNA directs the
protein's formation

Introduction

class = "introduction" (a) Nasal sinus cells (viewed with a light microscope), (b) onion cells (viewed with a light microscope), and (c) *Vibrio tasmaniensis* bacterial cells (seen through a scanning electron microscope) are from very different organisms, yet all share certain basic cell structure characteristics. (credit a: modification of work by Ed Uthman, MD; credit b: modification of work by Umberto Salvagnin; credit c: modification of work by Anthony D'Onofrio, William H. Fowle, Eric J. Stewart, and Kim Lewis of the Lewis Lab at Northeastern University; scale-bar data from Matt Russell)



Close your eyes and picture a brick wall. What is the wall's basic building block? It is a single brick. Like a brick wall, cells are the building blocks that make up your body.

Your body has many kinds of cells, each specialized for a specific purpose. Just as we use a variety of materials to build a home, the human body is constructed from many cell types. For example, epithelial cells protect the body's surface and cover

the organs and body cavities within. Bone cells help to support and protect the body. Immune system cells fight invading bacteria. Additionally, blood and blood cells carry nutrients and oxygen throughout the body while removing carbon dioxide. Each of these cell types plays a vital role during the body's growth, development, and day-to-day maintenance. In spite of their enormous variety, however, cells from all organisms—even ones as diverse as bacteria, onion, and human—share certain fundamental characteristics.

Studying Cells

By the end of this section, you will be able to do the following:

- Describe the role of cells in organisms
- Compare and contrast light microscopy and electron microscopy
- Summarize cell theory

A cell is the smallest unit of a living thing. Whether comprised of one cell (like bacteria) or many cells (like a human), we call it an organism. Thus, cells are the basic building blocks of all organisms.

Several cells of one kind that interconnect with each other and perform a shared function form tissues. These tissues combine to form an organ (your stomach, heart, or brain), and several organs comprise an organ system (such as the digestive system, circulatory system, or nervous system). Several systems that function together form an organism (like a human being). Here, we will examine the structure and function of cells.

There are many types of cells, which scientists group into one of two broad categories: prokaryotic and eukaryotic. For example, we classify both animal and plant cells as eukaryotic cells; whereas, we classify bacterial cells as prokaryotic. Before discussing the criteria for determining whether a cell is prokaryotic or eukaryotic, we will first

examine how biologists study cells.

(a) Most light microscopes in a college biology lab can magnify cells up to approximately 400 times and have a resolution of about 200 nanometers. (b) Electron microscopes provide a much higher magnification, 100,000x, and have a resolution of 50 picometers. (credit a: modification of work by "GcG"/Wikimedia Commons; credit b: modification of work by Evan Bench) (a) These *Salmonella* bacteria appear as tiny purple dots when viewed with a light microscope. (b) This scanning electron microscope micrograph shows *Salmonella* bacteria (in red) invading human cells (yellow). Even though subfigure (b) shows a different *Salmonella* specimen than subfigure (a), you can still observe the comparative increase in magnification and detail. (credit a: modification of work by CDC/Armed Forces Institute of Pathology, Charles N. Farmer, Rocky Mountain Laboratories; credit b: modification of work by NIAID, NIH; scale-bar data from Matt Russell)

Microscopy

Cells vary in size. With few exceptions, we cannot see individual cells with the naked eye, so scientists use microscopes (micro- = “small”; -scope = “to look at”) to study them. A **microscope** is an instrument that magnifies an object. We photograph most cells with a microscope, so we can call these images micrographs.

The optics of a microscope's lenses change the image orientation that the user sees. A specimen that is right-side up and facing right on the microscope slide will appear upside-down and facing left when one views through a microscope, and vice versa. Similarly, if one moves the slide left while looking through the microscope, it will appear to move right, and if one moves it down, it will seem to move up. This occurs because microscopes use two sets of lenses to magnify the image. Because of the manner by which light travels through the lenses, this two lens system produces an inverted image (binocular, or dissecting microscopes, work in a similar manner, but include an additional magnification system that makes the final image appear to be upright).

Light Microscopes

To give you a sense of cell size, a typical human red blood cell is about eight millionths of a meter or eight micrometers (abbreviated as eight μm) in diameter. A pin head is about two thousandths of a meter (two mm) in diameter. That means about 250 red blood cells could fit on a pinhead.

Most student microscopes are **light microscopes** ([\[link\]](#)a). Visible light passes and bends through the lens system to enable the user to see the specimen. Light microscopes are advantageous for viewing living organisms, but since individual cells are

generally transparent, their components are not distinguishable unless they are colored with special stains. Staining, however, usually kills the cells.

Light microscopes that undergraduates commonly use in the laboratory magnify up to approximately 400 times. Two parameters that are important in microscopy are magnification and resolving power. Magnification is the process of enlarging an object in appearance. Resolving power is the microscope's ability to distinguish two adjacent structures as separate: the higher the resolution, the better the image's clarity and detail. When one uses oil immersion lenses to study small objects, magnification usually increases to 1,000 times. In order to gain a better understanding of cellular structure and function, scientists typically use electron microscopes.



(a)



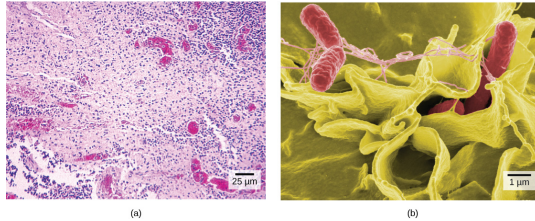
(b)

Electron Microscopes

In contrast to light microscopes, **electron microscopes** ([link](#)**b**) use a beam of electrons instead of a beam of light. Not only does this allow for higher magnification and, thus, more detail ([link](#)), it also provides higher resolving power. The method to prepare the specimen for viewing with an electron microscope kills the specimen. Electrons have short wavelengths (shorter than photons) that move best in a vacuum, so we cannot view living cells with an electron microscope.

In a scanning electron microscope, a beam of electrons moves back and forth across a cell's

surface, creating details of cell surface characteristics. In a transmission electron microscope, the electron beam penetrates the cell and provides details of a cell's internal structures. As you might imagine, electron microscopes are significantly more bulky and expensive than light microscopes.



Link to Learning

For another perspective on cell size, try the HowBig interactive at [this site](#).

Cell Theory

The microscopes we use today are far more complex than those that Dutch shopkeeper Antony van Leeuwenhoek, used in the 1600s. Skilled in crafting lenses, van Leeuwenhoek observed the movements of single-celled organisms, which he collectively termed “animalcules.”

In the 1665 publication *Micrographia*, experimental scientist Robert Hooke coined the term “cell” for the box-like structures he observed when viewing cork tissue through a lens. In the 1670s, van Leeuwenhoek discovered bacteria and protozoa. Later advances in lenses, microscope construction, and staining techniques enabled other scientists to see some components inside cells.

By the late 1830s, botanist Matthias Schleiden and zoologist Theodor Schwann were studying tissues and proposed the **unified cell theory**, which states that one or more cells comprise all living things, the cell is the basic unit of life, and new cells arise from existing cells. Rudolf Virchow later made important contributions to this theory.

Career Connection

Cytotechnologist

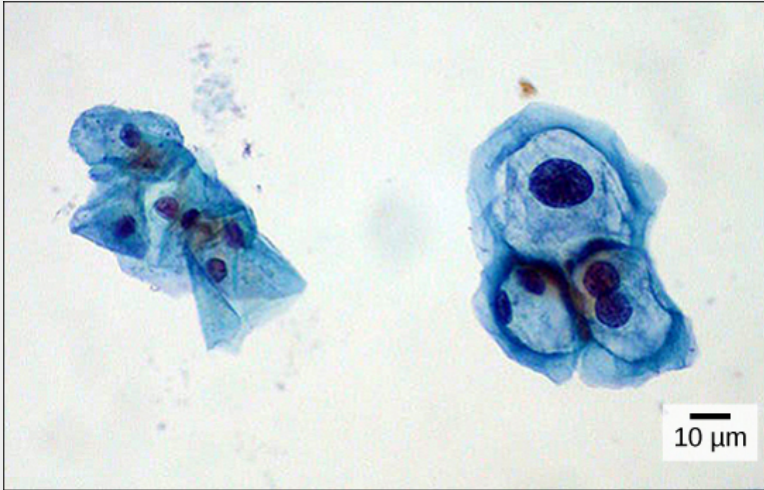
Have you ever heard of a medical test called a Pap smear ([\[link\]](#))? In this test, a doctor takes a small sample of cells from the patient's uterine cervix and sends it to a medical lab where a cytotechnologist stains the cells and examines them for any changes that could indicate cervical cancer or a microbial infection.

Cytotechnologists (cyto- = “cell”) are professionals who study cells via microscopic examinations and other laboratory tests. They are trained to

determine which cellular changes are within normal limits and which are abnormal. Their focus is not limited to cervical cells. They study cellular specimens that come from all organs. When they notice abnormalities, they consult a pathologist, a medical doctor who interprets and diagnoses changes that disease in body tissue and fluids cause.

Cytotechnologists play a vital role in saving people's lives. When doctors discover abnormalities early, a patient's treatment can begin sooner, which usually increases the chances of a successful outcome.

These uterine cervix cells, viewed through a light microscope, are from a Pap smear. Normal cells are on the left. The cells on the right are infected with human papillomavirus (HPV). Notice that the infected cells are larger. Also, two of these cells each have two nuclei instead of one, the normal number. (credit: modification of work by Ed Uthman, MD; scale-bar data from Matt Russell)



Section Summary

A cell is the smallest unit of life. Most cells are so tiny that we cannot see them with the naked eye. Therefore, scientists use microscopes to study cells. Electron microscopes provide higher magnification, higher resolution, and more detail than light microscopes. The unified cell theory states that one or more cells comprise all organisms, the cell is the basic unit of life, and new cells arise from existing cells.

Review Questions

When viewing a specimen through a light microscope, scientists use _____ to distinguish the individual components of cells.

1. a beam of electrons
2. radioactive isotopes
3. special stains
4. high temperatures

C

The _____ is the basic unit of life.

1. organism
2. cell
3. tissue
4. organ

B

Critical Thinking Questions

In your everyday life, you have probably noticed that certain instruments are ideal for

certain situations. For example, you would use a spoon rather than a fork to eat soup because a spoon is shaped for scooping, while soup would slip between the tines of a fork. The use of ideal instruments also applies in science. In what situation(s) would the use of a light microscope be ideal, and why?

A light microscope would be ideal when viewing a small living organism, especially when the cell has been stained to reveal details.

In what situation(s) would the use of a scanning electron microscope be ideal, and why?

A scanning electron microscope would be ideal when you want to view the minute details of a cell's surface, because its beam of electrons moves back and forth over the surface to convey the image.

In what situation(s) would a transmission electron microscope be ideal, and why?

A transmission electron microscope would be ideal for viewing the cell's internal structures, because many of the internal structures have

membranes that are not visible by the light microscope.

What are the advantages and disadvantages of each of these types of microscopes?

The advantages of light microscopes are that they are easily obtained, and the light beam does not kill the cells. However, typical light microscopes are somewhat limited in the amount of detail they can reveal. Electron microscopes are ideal because you can view intricate details, but they are bulky and costly, and preparation for the microscopic examination kills the specimen.

Explain how the formation of an adult human follows the cell theory.

The cell theory states:

1. All living things are made of cells.;
2. Cells are the most basic unit of life.;
3. New cells arise from existing cells.

All humans are multicellular organisms whose smallest building blocks are cells. Adult humans begin with the fusion of a male gamete cell

with a female gamete cell to form a fertilized egg (single cell). That cell then divides into two cells, which each divides into two more cells, and so forth until all the cells of a human embryo are made. As the embryo passes through all the developmental stages to make an adult human, the cells that are added arise from division of existing cells.

Glossary

cell theory

see unified cell theory

electron microscope

an instrument that magnifies an object using an electron beam that passes and bends through a lens system to visualize a specimen

light microscope

an instrument that magnifies an object using a beam of visible light that passes and bends through a lens system to visualize a specimen

microscope

an instrument that magnifies an object

unified cell theory

a biological concept that states that one or more cells comprise all organisms; the cell is the basic unit of life; and new cells arise from

existing cells

Prokaryotic Cells

By the end of this section, you will be able to do the following:

- Name examples of prokaryotic and eukaryotic organisms
- Compare and contrast prokaryotic and eukaryotic cells
- Describe the relative sizes of different cells
- Explain why cells must be small

Cells fall into one of two broad categories: prokaryotic and eukaryotic. We classify only the predominantly single-celled organisms Bacteria and Archaea as prokaryotes (pro- = “before”; -kary- = “nucleus”). Animal cells, plants, fungi, and protists are all eukaryotes (eu- = “true”).

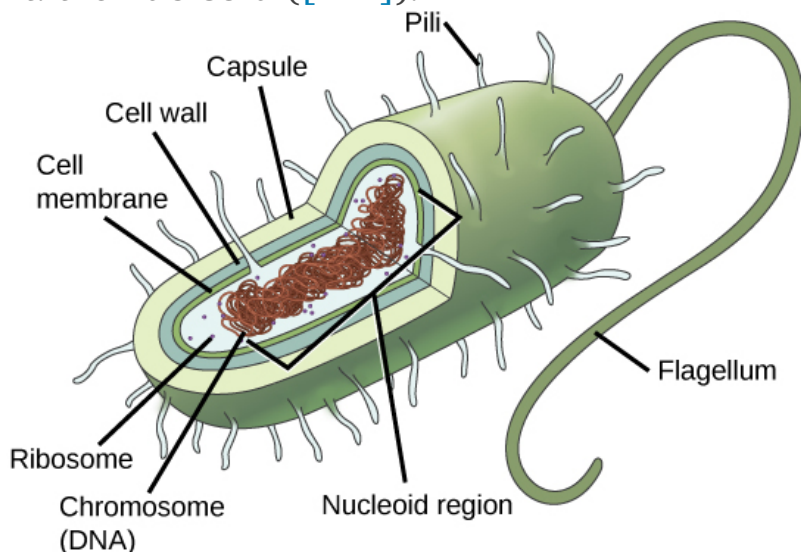
This figure shows the generalized structure of a prokaryotic cell. All prokaryotes have chromosomal DNA localized in a nucleoid, ribosomes, a cell membrane, and a cell wall. The other structures shown are present in some, but not all, bacteria. This figure shows relative sizes of microbes on a logarithmic scale (recall that each unit of increase in a logarithmic scale represents a 10-fold increase in the quantity measured).

Components of Prokaryotic Cells

All cells share four common components: 1) a

plasma membrane, an outer covering that separates the cell's interior from its surrounding environment; 2) cytoplasm, consisting of a jelly-like cytosol within the cell in which there are other cellular components; 3) DNA, the cell's genetic material; and 4) ribosomes, which synthesize proteins. However, prokaryotes differ from eukaryotic cells in several ways.

A **prokaryote** is a simple, mostly single-celled (unicellular) organism that lacks a nucleus, or any other membrane-bound organelle. We will shortly come to see that this is significantly different in eukaryotes. Prokaryotic DNA is in the cell's central part: the **nucleoid** ([\[link\]](#)).



Most prokaryotes have a peptidoglycan cell wall and many have a polysaccharide capsule ([\[link\]](#)). The cell wall acts as an extra layer of protection, helps

the cell maintain its shape, and prevents dehydration. The capsule enables the cell to attach to surfaces in its environment. Some prokaryotes have flagella, pili, or fimbriae. Flagella are used for locomotion. Pili exchange genetic material during conjugation, the process by which one bacterium transfers genetic material to another through direct contact. Bacteria use fimbriae to attach to a host cell.

Career Connection

Microbiologist

The most effective action anyone can take to prevent the spread of contagious illnesses is to wash his or her hands. Why? Because microbes (organisms so tiny that they can only be seen with microscopes) are ubiquitous. They live on doorknobs, money, your hands, and many other surfaces. If someone sneezes into his hand and touches a doorknob, and afterwards you touch that same doorknob, the microbes from the sneezer's mucus are now on your hands. If you touch your hands to your mouth, nose, or eyes, those microbes can enter your body and could make you sick. However, not all microbes (also called microorganisms) cause disease; most are actually beneficial. You have microbes in your gut that make vitamin K. Other microorganisms are used to ferment beer and wine.

Microbiologists are scientists who study microbes. Microbiologists can pursue a number of careers. Not only do they work in the food industry, they are also employed in the veterinary and medical fields. They can work in the pharmaceutical sector, serving key roles in research and development by identifying new antibiotic sources that can treat bacterial infections.

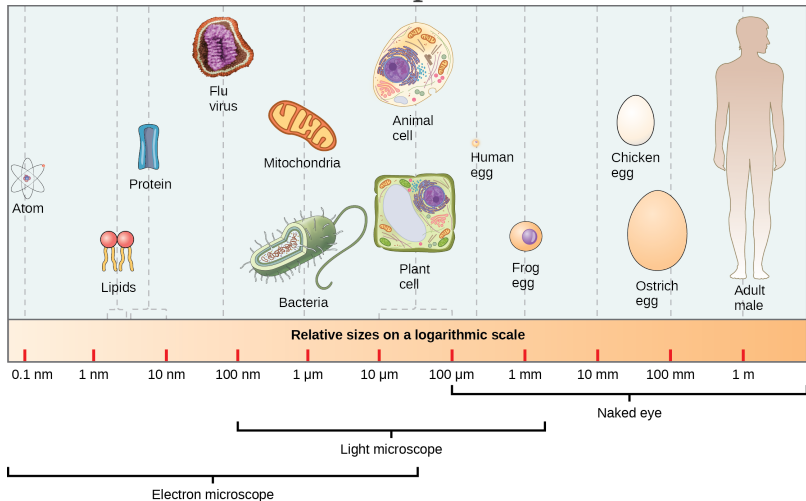
Environmental microbiologists may look for new ways to use specially selected or genetically engineered microbes to remove pollutants from soil or groundwater, as well as hazardous elements from contaminated sites. We call using these microbes bioremediation technologies.

Microbiologists can also work in the bioinformatics field, providing specialized knowledge and insight for designing, developing, and specificity of computer models of, for example, bacterial epidemics.

Cell Size

At 0.1 to 5.0 μm in diameter, prokaryotic cells are significantly smaller than eukaryotic cells, which have diameters ranging from 10 to 100 μm ([\[link\]](#)). The prokaryotes' small size allows ions and organic molecules that enter them to quickly diffuse to other parts of the cell. Similarly, any wastes produced within a prokaryotic cell can quickly diffuse. This is

not the case in eukaryotic cells, which have developed different structural adaptations to enhance intracellular transport.

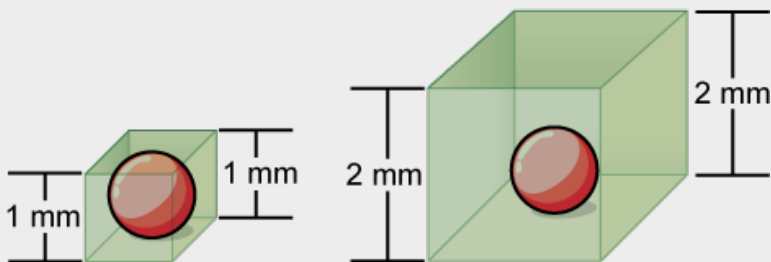


Small size, in general, is necessary for all cells, whether prokaryotic or eukaryotic. Let's examine why that is so. First, we'll consider the area and volume of a typical cell. Not all cells are spherical in shape, but most tend to approximate a sphere. You may remember from your high school geometry course that the formula for the surface area of a sphere is $4\pi r^2$, while the formula for its volume is $\frac{4\pi r^3}{3}$. Thus, as the radius of a cell increases, its surface area increases as the square of its radius, but its volume increases as the cube of its radius (much more rapidly). Therefore, as a cell increases in size, its surface area-to-volume ratio decreases. This same principle would apply if the cell had a cube shape ([\[link\]](#)). If the cell grows too large, the plasma membrane will not have sufficient surface area to

support the rate of diffusion required for the increased volume. In other words, as a cell grows, it becomes less efficient. One way to become more efficient is to divide. Another way is to develop organelles that perform specific tasks. These adaptations lead to developing more sophisticated cells, which we call eukaryotic cells.

Visual Connection

Notice that as a cell increases in size, its surface area-to-volume ratio decreases. When there is insufficient surface area to support a cell's increasing volume, a cell will either divide or die. The cell on the left has a volume of 1 mm^3 and a surface area of 6 mm^2 , with a surface area-to-volume ratio of 6 to 1; whereas, the cell on the right has a volume of 8 mm^3 and a surface area of 24 mm^2 , with a surface area-to-volume ratio of 3 to 1.



Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

Section Summary

Prokaryotes are single-celled organisms of the domains Bacteria and Archaea. All prokaryotes have plasma membranes, cytoplasm, ribosomes, and DNA that is not membrane-bound. Most have peptidoglycan cell walls and many have polysaccharide capsules. Prokaryotic cells range in diameter from 0.1 to 5.0 μm .

As a cell increases in size, its surface area-to-volume ratio decreases. If the cell grows too large, the plasma membrane will not have sufficient surface area to support the rate of diffusion required for the increased volume.

Visual Connection Questions

[\[link\]](#) Prokaryotic cells are much smaller than eukaryotic cells. What advantages might small cell size confer on a cell? What advantages might large cell size have?

[\[link\]](#) Substances can diffuse more quickly through small cells. Small cells have no need

for organelles and therefore do not need to expend energy getting substances across organelle membranes. Large cells have organelles that can separate cellular processes, enabling them to build molecules that are more complex.

Review Questions

Prokaryotes depend on _____ to obtain some materials and to get rid of wastes.

1. ribosomes
2. flagella
3. cell division
4. diffusion

D

Bacteria that lack fimbriae are less likely to _____.

1. adhere to cell surfaces
2. swim through bodily fluids
3. synthesize proteins
4. retain the ability to divide

A

Which of the following organisms is a prokaryote?

1. amoeba
2. influenza A virus
3. charophyte algae
4. *E. coli*

D

Critical Thinking Questions

Antibiotics are medicines that are used to fight bacterial infections. These medicines kill prokaryotic cells without harming human cells. What part or parts of the bacterial cell do you think antibiotics target? Why?

The cell wall would be targeted by antibiotics as well as the bacteria's ability to replicate. This would inhibit the bacteria's ability to reproduce, and it would compromise its defense mechanisms.

Explain why not all microbes are harmful.

Some microbes are beneficial. For instance, *E. coli* bacteria populate the human gut and help break down fiber in the diet. Some foods such as yogurt are formed by bacteria.

Glossary

nucleoid

central part of a prokaryotic cell's central part where the chromosome is located

prokaryote

unicellular organism that lacks a nucleus or any other membrane-bound organelle

Eukaryotic Cells

By the end of this section, you will be able to do the following:

- Describe the structure of eukaryotic cells
- Compare animal cells with plant cells
- State the role of the plasma membrane
- Summarize the functions of the major cell organelles

Have you ever heard the phrase “form follows function?” It’s a philosophy that many industries follow. In architecture, this means that buildings should be constructed to support the activities that will be carried out inside them. For example, a skyscraper should include several elevator banks. A hospital should have its emergency room easily accessible.

Our natural world also utilizes the principle of form following function, especially in cell biology, and this will become clear as we explore eukaryotic cells ([\[link\]](#)). Unlike prokaryotic cells, **eukaryotic cells** have: 1) a membrane-bound nucleus; 2) numerous membrane-bound **organelles** such as the endoplasmic reticulum, Golgi apparatus, chloroplasts, mitochondria, and others; and 3) several, rod-shaped chromosomes. Because a membrane surrounds eukaryotic cell’s nucleus, it has a “true nucleus.” The word “organelle” means “little organ,” and, as we already mentioned,

organelles have specialized cellular functions, just as your body's organs have specialized functions.

At this point, it should be clear to you that eukaryotic cells have a more complex structure than prokaryotic cells. Organelles allow different functions to be compartmentalized in different areas of the cell. Before turning to organelles, let's first examine two important components of the cell: the plasma membrane and the cytoplasm.

Visual Connection

These figures show the major organelles and other cell components of (a) a typical animal cell and (b) a typical eukaryotic plant cell. The plant cell has a cell wall, chloroplasts, plastids, and a central vacuole—structures not in animal cells. Most cells do not have lysosomes or centrosomes.

Nucleus

Nuclear envelope:

membrane enclosing the nucleus. Protein-lined pores allow material to move in and out.

Chromatin: DNA plus associated proteins.

Nucleolus:

condensed region where ribosomes are formed.

Peroxisome:

metabolizes waste

Endoplasmic reticulum

Rough: associated with ribosomes; makes secretory and membrane proteins.

Smooth: makes lipids.

Plasmodesmata

channels connect two plant cells

Cell wall maintains cell shape

Plasma membrane

Cytoplasm

Central Vacuole

filled with cell sap that maintains pressure against cell wall

Cytoskeleton

microtubules
intermediate filaments
microfilaments

Chloroplast site of photosynthesis

Plastid store pigments

Cytoskeleton

Microtubules: form the mitotic spindle and maintain cell shape.

Centrosome: microtubule-organizing center.

Intermediate filaments: fibrous proteins that hold organelles in place.

Microfilaments:

fibrous proteins; form the cellular cortex.

Plasma membrane

Lysosome: digests food and waste materials.

Golgi apparatus: modifies proteins.

Cytoplasm

Vacuole

Mitochondria: produce energy.

Endoplasmic Reticulum

smooth

rough

Nucleus contains chromatin, a nuclear envelope, and a nucleolus, as in an animal cell

Ribosomes

Golgi apparatus

Mitochondria

Peroxisome

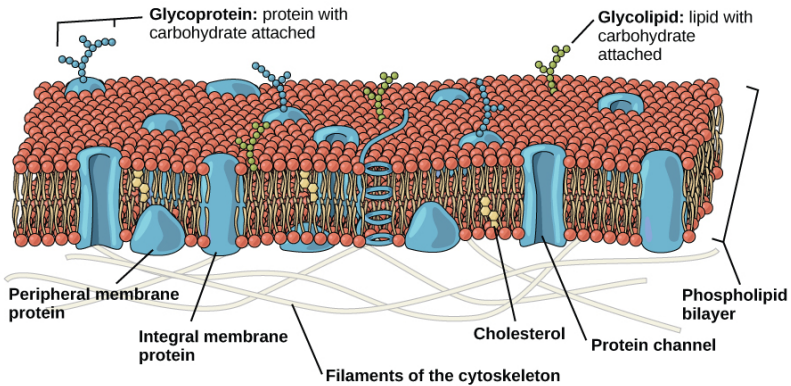
If the nucleolus were not able to carry out its function, what other cellular organelles would be

affected?

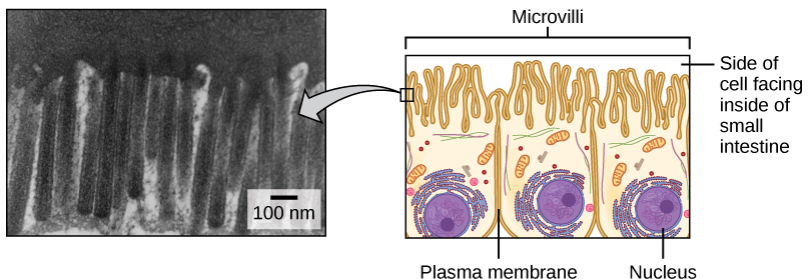
The eukaryotic plasma membrane is a phospholipid bilayer with proteins and cholesterol embedded in it. Microvilli, as they appear on cells lining the small intestine, increase the surface area available for absorption. These microvilli are only on the area of the plasma membrane that faces the cavity from which substances will be absorbed. (credit "micrograph": modification of work by Louisa Howard)

The Plasma Membrane

Like prokaryotes, eukaryotic cells have a **plasma membrane** ([link](#)), a phospholipid bilayer with embedded proteins that separates the internal contents of the cell from its surrounding environment. A phospholipid is a lipid molecule with two fatty acid chains and a phosphate-containing group. The plasma membrane controls the passage of organic molecules, ions, water, and oxygen into and out of the cell. Wastes (such as carbon dioxide and ammonia) also leave the cell by passing through the plasma membrane.



The plasma membranes of cells that specialize in absorption fold into fingerlike projections that we call microvilli (singular = microvillus); ([\[link\]](#)). Such cells typically line the small intestine, the organ that absorbs nutrients from digested food. This is an excellent example of form following function. People with celiac disease have an immune response to gluten, which is a protein in wheat, barley, and rye. The immune response damages microvilli, and thus, afflicted individuals cannot absorb nutrients. This leads to malnutrition, cramping, and diarrhea. Patients suffering from celiac disease must follow a gluten-free diet.



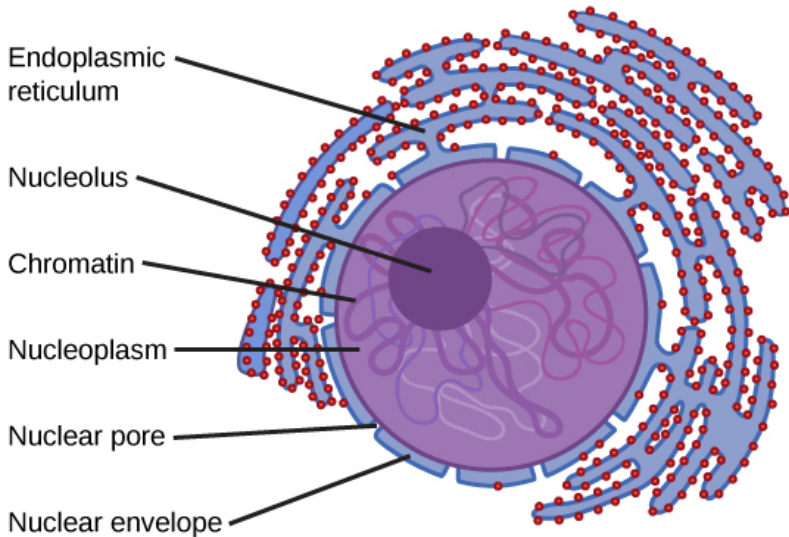
The Cytoplasm

The **cytoplasm** is the cell's entire region between the plasma membrane and the nuclear envelope (a structure we will discuss shortly). It is comprised of organelles suspended in the gel-like **cytosol**, the cytoskeleton, and various chemicals ([\[link\]](#)). Even though the cytoplasm consists of 70 to 80 percent water, it has a semi-solid consistency, which comes from the proteins within it. However, proteins are not the only organic molecules in the cytoplasm. Glucose and other simple sugars, polysaccharides, amino acids, nucleic acids, fatty acids, and derivatives of glycerol are also there. Ions of sodium, potassium, calcium, and many other elements also dissolve in the cytoplasm. Many metabolic reactions, including protein synthesis, take place in the cytoplasm.

The nucleus stores chromatin (DNA plus proteins) in a gel-like substance called the nucleoplasm. The nucleolus is a condensed chromatin region where ribosome synthesis occurs. We call the nucleus' boundary the nuclear envelope. It consists of two phospholipid bilayers: an outer and an inner membrane. The nuclear membrane is continuous with the endoplasmic reticulum. Nuclear pores allow substances to enter and exit the nucleus. (a) This image shows various levels of chromatin's organization (DNA and protein). (b) This image shows paired chromosomes. (credit b: modification of work by NIH; scale-bar data from Matt Russell)

The Nucleus

Typically, the nucleus is the most prominent organelle in a cell ([\[link\]](#)). The **nucleus** (plural = nuclei) houses the cell's DNA and directs the synthesis of ribosomes and proteins. Let's look at it in more detail ([\[link\]](#)).



The Nuclear Envelope

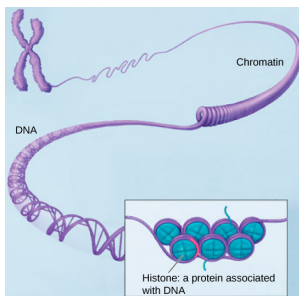
The **nuclear envelope** is a double-membrane structure that constitutes the nucleus' outermost portion ([\[link\]](#)). Both the nuclear envelope's inner and outer membranes are phospholipid bilayers.

The nuclear envelope is punctuated with pores that control the passage of ions, molecules, and RNA between the nucleoplasm and cytoplasm. The **nucleoplasm** is the semi-solid fluid inside the

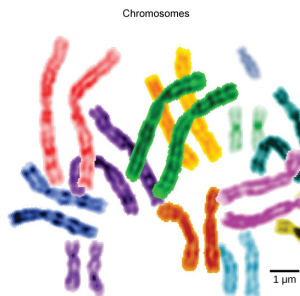
nucleus, where we find the chromatin and the nucleolus.

Chromatin and Chromosomes

To understand chromatin, it is helpful to first explore **chromosomes**, structures within the nucleus that are made up of DNA, the hereditary material. You may remember that in prokaryotes, DNA is organized into a single circular chromosome. In eukaryotes, chromosomes are linear structures. Every eukaryotic species has a specific number of chromosomes in the nucleus of each cell. For example, in humans, the chromosome number is 46, while in fruit flies, it is eight. Chromosomes are only visible and distinguishable from one another when the cell is getting ready to divide. When the cell is in the growth and maintenance phases of its life cycle, proteins attach to chromosomes, and they resemble an unwound, jumbled bunch of threads. We call these unwound protein-chromosome complexes **chromatin** ([\[link\]](#)). Chromatin describes the material that makes up the chromosomes both when condensed and decondensed.



(a)



(b)

The Nucleolus

We already know that the nucleus directs the synthesis of ribosomes, but how does it do this? Some chromosomes have sections of DNA that encode ribosomal RNA. A darkly staining area within the nucleus called the **nucleolus** (plural = nucleoli) aggregates the ribosomal RNA with associated proteins to assemble the ribosomal subunits that are then transported out through the pores in the nuclear envelope to the cytoplasm.

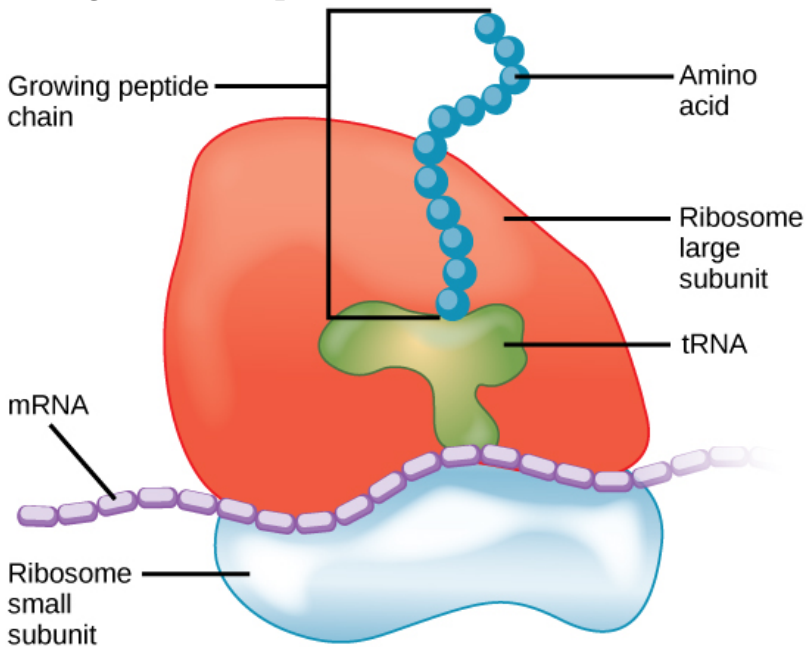
A large subunit (top) and a small subunit (bottom) comprise ribosomes. During protein synthesis, ribosomes assemble amino acids into proteins.

Ribosomes

Ribosomes are the cellular structures responsible for protein synthesis. When we view them through an electron microscope, ribosomes appear either as clusters (polyribosomes) or single, tiny dots that float freely in the cytoplasm. They may be attached to the plasma membrane's cytoplasmic side or the endoplasmic reticulum's cytoplasmic side and the nuclear envelope's outer membrane ([\[link\]](#)).

Electron microscopy shows us that ribosomes, which are large protein and RNA complexes, consist of two subunits, large and small ([\[link\]](#)). Ribosomes receive their “orders” for protein synthesis from the nucleus where the DNA transcribes into messenger RNA (mRNA). The mRNA travels to the ribosomes, which

translate the code provided by the sequence of the nitrogenous bases in the mRNA into a specific order of amino acids in a protein. Amino acids are the building blocks of proteins.



Because protein synthesis is an essential function of all cells (including enzymes, hormones, antibodies, pigments, structural components, and surface receptors), there are ribosomes in practically every cell. Ribosomes are particularly abundant in cells that synthesize large amounts of protein. For example, the pancreas is responsible for creating several digestive enzymes and the cells that produce these enzymes contain many ribosomes. Thus, we see another example of form following function.

This electron micrograph shows a mitochondrion through an electron microscope. This organelle has

an outer membrane and an inner membrane. The inner membrane contains folds, called cristae, which increase its surface area. We call the space between the two membranes the intermembrane space, and the space inside the inner membrane the mitochondrial matrix. ATP synthesis takes place on the inner membrane. (credit: modification of work by Matthew Britton; scale-bar data from Matt Russell)

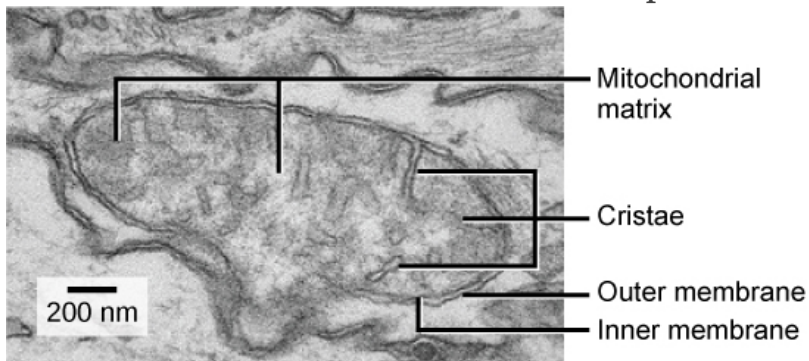
Mitochondria

Scientists often call **mitochondria** (singular = mitochondrion) the cell's “powerhouses” or “energy factories” because they are responsible for making adenosine triphosphate (ATP), the cell's main energy-carrying molecule. ATP represents the cell's short-term stored energy. Cellular respiration is the process of making ATP using the chemical energy in glucose and other nutrients. In mitochondria, this process uses oxygen and produces carbon dioxide as a waste product. In fact, the carbon dioxide that you exhale with every breath comes from the cellular reactions that produce carbon dioxide as a byproduct.

In keeping with our theme of form following function, it is important to point out that muscle cells have a very high concentration of mitochondria that produce ATP. Your muscle cells need considerable energy to keep your body

moving. When your cells don't get enough oxygen, they do not make much ATP. Instead, producing lactic acid accompanies the small amount of ATP they make in the absence of oxygen.

Mitochondria are oval-shaped, double membrane organelles ([\[link\]](#)) that have their own ribosomes and DNA. Each membrane is a phospholipid bilayer embedded with proteins. The inner layer has folds called cristae. We call the area surrounded by the folds the mitochondrial matrix. The cristae and the matrix have different roles in cellular respiration.



Peroxisomes

Peroxisomes are small, round organelles enclosed by single membranes. They carry out oxidation reactions that break down fatty acids and amino acids. They also detoxify many poisons that may enter the body. (Many of these oxidation reactions release hydrogen peroxide, H_2O_2 , which would be damaging to cells; however, when these reactions

are confined to peroxisomes, enzymes safely break down the H_2O_2 into oxygen and water.) For example, peroxisomes in liver cells detoxify alcohol. Glyoxysomes, which are specialized peroxisomes in plants, are responsible for converting stored fats into sugars. Plant cells contain many different types of peroxisomes that play a role in metabolism, pathogene defense, and stress response, to mention a few.

Vesicles and Vacuoles

Vesicles and **vacuoles** are membrane-bound sacs that function in storage and transport. Other than the fact that vacuoles are somewhat larger than vesicles, there is a very subtle distinction between them. Vesicle membranes can fuse with either the plasma membrane or other membrane systems within the cell. Additionally, some agents such as enzymes within plant vacuoles break down macromolecules. The vacuole's membrane does not fuse with the membranes of other cellular components.

The centrosome consists of two centrioles that lie at right angles to each other. Each centriole is a cylinder comprised of nine triplets of microtubules. Nontubulin proteins (indicated by the green lines) hold the microtubule triplets together. Cellulose is a long chain of β -glucose molecules connected by a 1-4 linkage. The dashed lines at each end of the

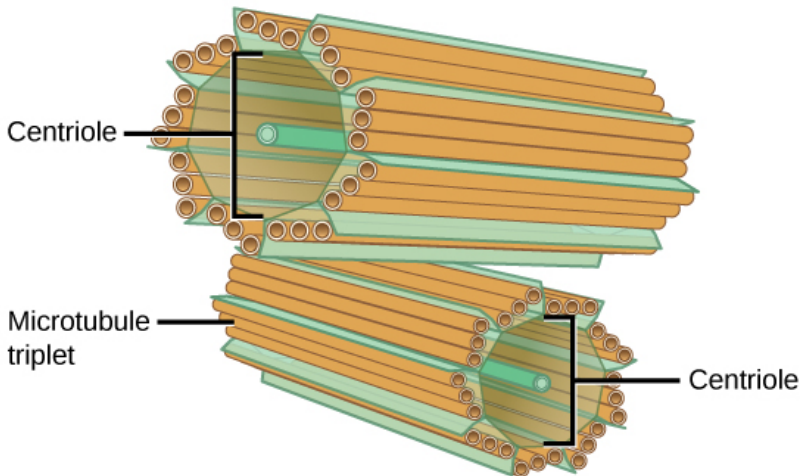
figure indicate a series of many more glucose units. The size of the page makes it impossible to portray an entire cellulose molecule. The chloroplast has an outer membrane, an inner membrane, and membrane structures - thylakoids that are stacked into grana. We call the space inside the thylakoid membranes the thylakoid space. The light harvesting reactions take place in the thylakoid membranes, and sugar synthesis takes place in the fluid inside the inner membrane, which we call the stroma. Chloroplasts also have their own genome, which is contained on a single circular chromosome.

Animal Cells versus Plant Cells

At this point, you know that each eukaryotic cell has a plasma membrane, cytoplasm, a nucleus, ribosomes, mitochondria, peroxisomes, and in some, vacuoles, but there are some striking differences between animal and plant cells. While both animal and plant cells have microtubule organizing centers (MTOCs), animal cells also have centrioles associated with the MTOC: a complex we call the centrosome. Animal cells each have a centrosome and lysosomes; whereas, most plant cells do not. Plant cells have a cell wall, chloroplasts and other specialized plastids, and a large central vacuole; whereas, animal cells do not.

The Centrosome

The **centrosome** is a microtubule-organizing center found near the nuclei of animal cells. It contains a pair of centrioles, two structures that lie perpendicular to each other ([\[link\]](#)). Each centriole is a cylinder of nine triplets of microtubules.



The centrosome (the organelle where all microtubules originate) replicates itself before a cell divides, and the centrioles appear to have some role in pulling the duplicated chromosomes to opposite ends of the dividing cell. However, the centriole's exact function in cell division isn't clear, because cells that have had the centrosome removed can still divide, and plant cells, which lack centrosomes, are capable of cell division.

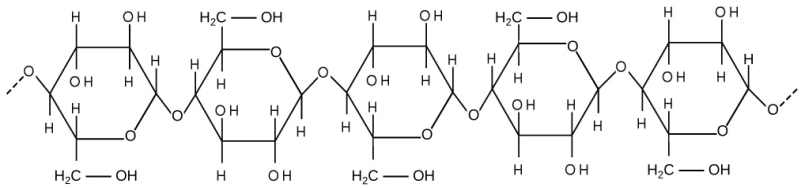
Lysosomes

Animal cells have another set of organelles that most plant cells do not: lysosomes. The **lysosomes**

are the cell's "garbage disposal." In plant cells, the digestive processes take place in vacuoles. Enzymes within the lysosomes aid in breaking down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles. These enzymes are active at a much lower pH than the cytoplasm's. Therefore, the pH within lysosomes is more acidic than the cytoplasm's pH. Many reactions that take place in the cytoplasm could not occur at a low pH, so again, the advantage of compartmentalizing the eukaryotic cell into organelles is apparent.

The Cell Wall

If you examine [\[link\]](#), the plant cell diagram, you will see a structure external to the plasma membrane. This is the **cell wall**, a rigid covering that protects the cell, provides structural support, and gives shape to the cell. Fungal and some protistan cells also have cell walls. While the prokaryotic cell walls' chief component is peptidoglycan, the major organic molecule in the plant (and some protists') cell wall is cellulose ([\[link\]](#)), a polysaccharide comprised of glucose units. Have you ever noticed that when you bite into a raw vegetable, like celery, it crunches? That's because you are tearing the celery cells' rigid cell walls with your teeth.

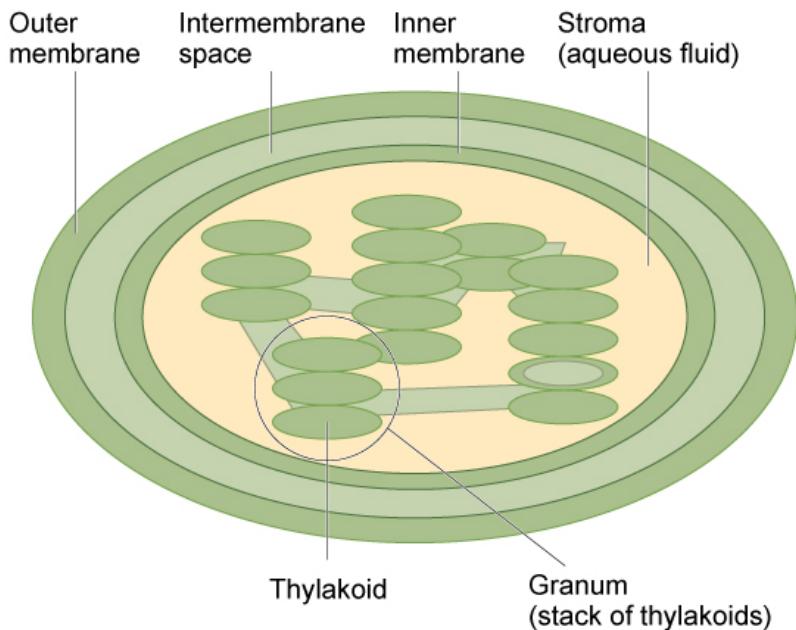


Chloroplasts

Like the mitochondria, chloroplasts have their own DNA and ribosomes, but chloroplasts have an entirely different function. **Chloroplasts** are plant cell organelles that carry out photosynthesis.

Photosynthesis is the series of reactions that use carbon dioxide, water, and light energy to make glucose and oxygen. This is a major difference between plants and animals. Plants (autotrophs) are able to make their own food, like sugars, while animals (heterotrophs) must ingest their food.

Like mitochondria, chloroplasts have outer and inner membranes, but within the space enclosed by a chloroplast's inner membrane is a set of interconnected and stacked fluid-filled membrane sacs we call thylakoids ([\[link\]](#)). Each thylakoid stack is a granum (plural = grana). We call the fluid enclosed by the inner membrane that surrounds the grana the stroma.



The chloroplasts contain a green pigment, **chlorophyll**, which captures the light energy that drives the reactions of photosynthesis. Like plant cells, photosynthetic protists also have chloroplasts. Some bacteria perform photosynthesis, but their chlorophyll is not relegated to an organelle.

Evolution Connection

Endosymbiosis

We have mentioned that both mitochondria and chloroplasts contain DNA and ribosomes. Have you wondered why? Strong evidence points to endosymbiosis as the explanation.

Symbiosis is a relationship in which organisms

from two separate species depend on each other for their survival. Endosymbiosis (endo- = “within”) is a mutually beneficial relationship in which one organism lives inside the other. Endosymbiotic relationships abound in nature. We have already mentioned that microbes that produce vitamin K live inside the human gut. This relationship is beneficial for us because we are unable to synthesize vitamin K. It is also beneficial for the microbes because they are protected from other organisms and from drying out, and they receive abundant food from the environment of the large intestine.

Scientists have long noticed that bacteria, mitochondria, and chloroplasts are similar in size. We also know that bacteria have DNA and ribosomes, just like mitochondria and chloroplasts. Scientists believe that host cells and bacteria formed an endosymbiotic relationship when the host cells ingested both aerobic and autotrophic bacteria (cyanobacteria) but did not destroy them. Through many millions of years of evolution, these ingested bacteria became more specialized in their functions, with the aerobic bacteria becoming mitochondria and the autotrophic bacteria becoming chloroplasts.

The Central Vacuole

Previously, we mentioned vacuoles as essential components of plant cells. If you look at [\[link\]](#)^b, you will see that plant cells each have a large central vacuole that occupies most of the cell's area. The **central vacuole** plays a key role in regulating the cell's concentration of water in changing environmental conditions. Have you ever noticed that if you forget to water a plant for a few days, it wilts? That's because as the water concentration in the soil becomes lower than the water concentration in the plant, water moves out of the central vacuoles and cytoplasm. As the central vacuole shrinks, it leaves the cell wall unsupported. This loss of support to the plant's cell walls results in the wilted appearance.

The central vacuole also supports the cell's expansion. When the central vacuole holds more water, the cell becomes larger without having to invest considerable energy in synthesizing new cytoplasm.

Section Summary

Like a prokaryotic cell, a eukaryotic cell has a plasma membrane, cytoplasm, and ribosomes, but a eukaryotic cell is typically larger than a prokaryotic cell, has a true nucleus (meaning a membrane surrounds its DNA), and has other membrane-bound organelles that allow for compartmentalizing

functions. The plasma membrane is a phospholipid bilayer embedded with proteins. The nucleus's nucleolus is the site of ribosome assembly. We find ribosomes either in the cytoplasm or attached to the cytoplasmic side of the plasma membrane or endoplasmic reticulum. They perform protein synthesis. Mitochondria participate in cellular respiration. They are responsible for the majority of ATP produced in the cell. Peroxisomes hydrolyze fatty acids, amino acids, and some toxins. Vesicles and vacuoles are storage and transport compartments. In plant cells, vacuoles also help break down macromolecules.

Animal cells also have a centrosome and lysosomes. The centrosome has two bodies perpendicular to each other, the centrioles, and has an unknown purpose in cell division. Lysosomes are the digestive organelles of animal cells.

Plant cells and plant-like cells each have a cell wall, chloroplasts, and a central vacuole. The plant cell wall, whose primary component is cellulose, protects the cell, provides structural support, and gives the cell shape. Photosynthesis takes place in chloroplasts. The central vacuole can expand without having to produce more cytoplasm.

Visual Connection Questions

[\[link\]](#) If the nucleolus were not able to carry out its function, what other cellular organelles would be affected?

[\[link\]](#) Free ribosomes and rough endoplasmic reticulum (which contains ribosomes) would not be able to form.

Review Questions

Which of the following is surrounded by two phospholipid bilayers?

1. the ribosomes
2. the vesicles
3. the cytoplasm
4. the nucleoplasm

D

Peroxisomes got their name because hydrogen peroxide is:

1. used in their detoxification reactions
2. produced during their oxidation reactions

3. incorporated into their membranes
 4. a cofactor for the organelles' enzymes
-

B

In plant cells, the function of the lysosomes is carried out by _____.

1. vacuoles
 2. peroxisomes
 3. ribosomes
 4. nuclei
-

A

Which of the following is both in eukaryotic and prokaryotic cells?

1. nucleus
 2. mitochondrion
 3. vacuole
 4. ribosomes
-

D

Tay-Sachs disease is a genetic disorder that

results in the destruction of neurons due to a buildup of sphingolipids in the cells. Which organelle is malfunctioning in Tay-Sachs?

1. lysosome
2. endoplasmic reticulum
3. peroxisome
4. mitochondria

A

Critical Thinking Questions

You already know that ribosomes are abundant in red blood cells. In what other cells of the body would you find them in great abundance? Why?

Ribosomes are abundant in muscle cells as well because muscle cells are constructed of the proteins made by the ribosomes.

What are the structural and functional similarities and differences between mitochondria and chloroplasts?

Both are similar in that they are enveloped in a double membrane, both have an intermembrane space, and both make ATP. Both mitochondria and chloroplasts have DNA, and mitochondria have inner folds called cristae and a matrix, while chloroplasts have chlorophyll and accessory pigments in the thylakoids that form stacks (grana) and a stroma.

Why are plasma membranes arranged as a bilayer rather than a monolayer?

The plasma membrane is a bilayer because the phospholipids that create it are amphiphilic (hydrophilic head, hydrophobic tail). If the plasma membrane was a monolayer, the hydrophobic tails of the phospholipids would be in direct contact with the inside of the cell. Since the cytoplasm is largely made of water, this interaction would not be stable, and would disrupt the plasma membrane of the cell as the tails were repulsed by the cytoplasm (in water, phospholipids spontaneously form spherical droplets with the hydrophilic heads facing outward to isolate the hydrophobic tails from the water). By having a bilayer, the hydrophilic heads are exposed to the aqueous cytoplasm and extracellular space, while the hydrophobic

tails interact with each other in the middle of the membrane.

Glossary

cell wall

rigid cell covering comprised of various molecules that protects the cell, provides structural support, and gives shape to the cell

central vacuole

large plant cell organelle that regulates the cell's storage compartment, holds water, and plays a significant role in cell growth as the site of macromolecule degradation

centrosome

region in animal cells made of two centrioles that serves as an organizing center for microtubules

chlorophyll

green pigment that captures the light energy that drives the light reactions of photosynthesis

chloroplast

plant cell organelle that carries out photosynthesis

chromatin

protein-DNA complex that serves as the chromosomes' building material

chromosome

structure within the nucleus that comprises chromatin that contains DNA, the hereditary material

cytoplasm

entire region between the plasma membrane and the nuclear envelope, consisting of organelles suspended in the gel-like cytosol, the cytoskeleton, and various chemicals

cytosol

the cytoplasm's gel-like material in which cell structures are suspended

eukaryotic cell

cell that has a membrane-bound nucleus and several other membrane-bound compartments or sacs

lysosome

organelle in an animal cell that functions as the cell's digestive component; it breaks down proteins, polysaccharides, lipids, nucleic acids, and even worn-out organelles

mitochondria

(singular = mitochondrion) cellular organelles responsible for carrying out

cellular respiration, resulting in producing ATP, the cell's main energy-carrying molecule

nuclear envelope

double-membrane structure that constitutes the nucleus' outermost portion

nucleolus

darkly staining body within the nucleus that is responsible for assembling ribosome subunits

nucleoplasm

semi-solid fluid inside the nucleus that contains the chromatin and nucleolus

nucleus

cell organelle that houses the cell's DNA and directs ribosome and protein synthesis

organelle

compartment or sac within a cell

peroxisome

small, round organelle that contains hydrogen peroxide, oxidizes fatty acids and amino acids, and detoxifies many poisons

plasma membrane

phospholipid bilayer with embedded (integral) or attached (peripheral) proteins, and separates the cell's internal content from

its surrounding environment

ribosome

cellular structure that carries out protein synthesis

vacuole

membrane-bound sac, somewhat larger than a vesicle, which functions in cellular storage and transport

vesicle

small, membrane-bound sac that functions in cellular storage and transport; its membrane is capable of fusing with the plasma membrane and the membranes of the endoplasmic reticulum and Golgi apparatus

The Endomembrane System and Proteins

By the end of this section, you will be able to do the following:

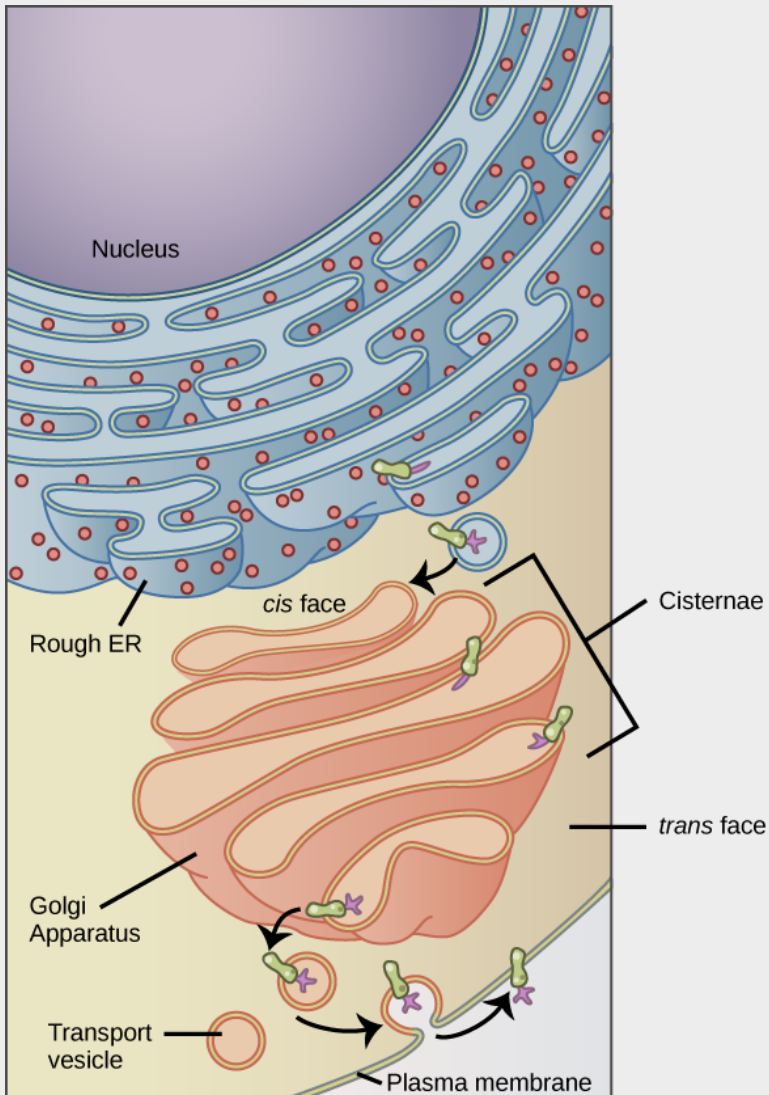
- List the components of the endomembrane system
- Recognize the relationship between the endomembrane system and its functions

The endomembrane system (endo = “within”) is a group of membranes and organelles ([\[link\]](#)) in eukaryotic cells that works together to modify, package, and transport lipids and proteins. It includes the nuclear envelope, lysosomes, and vesicles, which we have already mentioned, and the endoplasmic reticulum and Golgi apparatus, which we will cover shortly. Although not technically *within* the cell, the plasma membrane is included in the endomembrane system because, as you will see, it interacts with the other endomembranous organelles. The endomembrane system does not include either mitochondria or chloroplast membranes.

Visual Connection

Membrane and secretory proteins are synthesized in the rough endoplasmic reticulum (RER). The RER also sometimes modifies proteins. In this illustration, an attachment of a (purple)

carbohydrate modifies a (green) integral membrane protein in the ER. Vesicles with the integral protein bud from the ER and fuse with the Golgi apparatus' cis face. As the protein passes along the Golgi's cisternae, the addition of more carbohydrates further modifies it. After its synthesis is complete, it exits as an integral membrane protein of the vesicle that buds from the Golgi's **trans** face. When the vesicle fuses with the cell membrane, the protein becomes an integral portion of that cell membrane. (credit: modification of work by Magnus Manske)



If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

This transmission electron micrograph shows the

rough endoplasmic reticulum and other organelles in a pancreatic cell. (credit: modification of work by Louisa Howard)

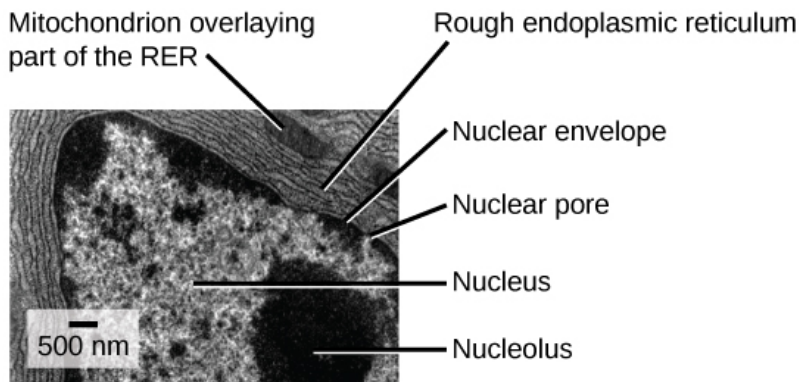
The Endoplasmic Reticulum

The **endoplasmic reticulum (ER)** ([\[link\]](#)) is a series of interconnected membranous sacs and tubules that collectively modifies proteins and synthesizes lipids. However, these two functions take place in separate areas of the ER: the rough ER and the smooth ER, respectively.

We call the ER tubules' hollow portion the lumen or cisternal space. The ER's membrane, which is a phospholipid bilayer embedded with proteins, is continuous with the nuclear envelope.

Rough ER

Scientists have named the **rough endoplasmic reticulum (RER)** as such because the ribosomes attached to its cytoplasmic surface give it a studded appearance when viewing it through an electron microscope ([\[link\]](#)).



Ribosomes transfer their newly synthesized proteins into the RER's lumen where they undergo structural modifications, such as folding or acquiring side chains. These modified proteins incorporate into cellular membranes—the ER or the ER's or other organelles' membranes. The proteins can also secrete from the cell (such as protein hormones, enzymes). The RER also makes phospholipids for cellular membranes.

If the phospholipids or modified proteins are not destined to stay in the RER, they will reach their destinations via transport vesicles that bud from the RER's membrane ([\[link\]](#)).

Since the RER is engaged in modifying proteins (such as enzymes, for example) that secrete from the cell, you would be correct in assuming that the RER is abundant in cells that secrete proteins. This is the case with liver cells, for example.

Smooth ER

The **smooth endoplasmic reticulum (SER)** is continuous with the RER but has few or no ribosomes on its cytoplasmic surface ([\[link\]](#)). SER functions include synthesis of carbohydrates, lipids, and steroid hormones; detoxification of medications and poisons; and storing calcium ions.

In muscle cells, a specialized SER, the sarcoplasmic reticulum, is responsible for storing calcium ions that are needed to trigger the muscle cells' coordinated contractions.

Link to Learning

You can watch an excellent animation of the endomembrane system [here](#). At the end of the animation, there is a short self-assessment.

Career Connection

Cardiologist

Heart disease is the leading cause of death in the United States. This is primarily due to our sedentary lifestyle and our high trans-fat diets. Heart failure is just one of many disabling heart conditions. Heart failure does not mean that the heart has stopped working. Rather, it means that the heart can't pump with sufficient force to transport oxygenated blood to all the vital organs.

Left untreated, heart failure can lead to kidney failure and other organ failure.

Cardiac muscle tissue comprises the heart's wall.

Heart failure occurs when cardiac muscle cells' endoplasmic reticula do not function properly. As a result, an insufficient number of calcium ions are available to trigger a sufficient contractile force.

Cardiologists (cardi- = “heart”; -ologist = “one who studies”) are doctors who specialize in treating heart diseases, including heart failure.

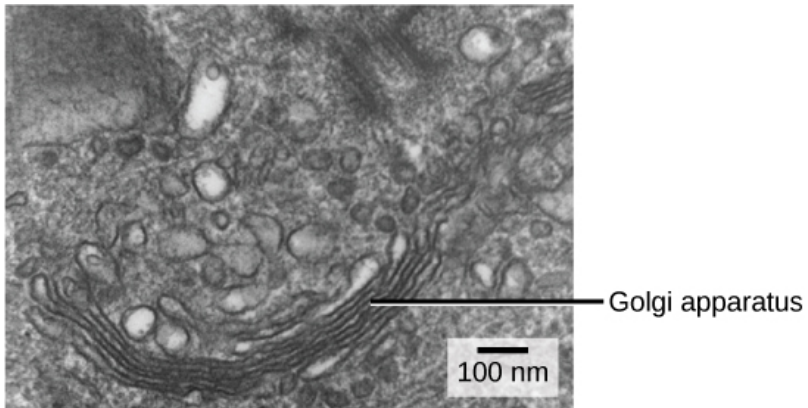
Cardiologists can diagnose heart failure via a physical examination, results from an electrocardiogram (ECG, a test that measures the heart's electrical activity), a chest X-ray to see whether the heart is enlarged, and other tests. If the cardiologist diagnoses heart failure, he or she will typically prescribe appropriate medications and recommend a reduced table salt intake and a supervised exercise program.

The Golgi apparatus in this white blood cell is visible as a stack of semicircular, flattened rings in the lower portion of the image. You can see several vesicles near the Golgi apparatus. (credit: modification of work by Louisa Howard)

The Golgi Apparatus

We have already mentioned that vesicles can bud

from the ER and transport their contents elsewhere, but where do the vesicles go? Before reaching their final destination, the lipids or proteins within the transport vesicles still need sorting, packaging, and tagging so that they end up in the right place. Sorting, tagging, packaging, and distributing lipids and proteins takes place in the **Golgi apparatus** (also called the Golgi body), a series of flattened membranes ([\[link\]](#)).



We call the Golgi apparatus' the *cis* face. The opposite side is the *trans* face. The transport vesicles that formed from the ER travel to the *cis* face, fuse with it, and empty their contents into the Golgi apparatus' lumen. As the proteins and lipids travel through the Golgi, they undergo further modifications that allow them to be sorted. The most frequent modification is adding short sugar molecule chains. These newly modified proteins and lipids then tag with phosphate groups or other small molecules in order to travel to their proper destinations.

Finally, the modified and tagged proteins are packaged into secretory vesicles that bud from the Golgi's *trans* face. While some of these vesicles deposit their contents into other cell parts where they will be used, other secretory vesicles fuse with the plasma membrane and release their contents outside the cell.

In another example of form following function, cells that engage in a great deal of secretory activity (such as salivary gland cells that secrete digestive enzymes or immune system cells that secrete antibodies) have an abundance of Golgi.

In plant cells, the Golgi apparatus has the additional role of synthesizing polysaccharides, some of which are incorporated into the cell wall and some of which other cell parts use.

Career Connection

Geneticist

Many diseases arise from genetic mutations that prevent synthesizing critical proteins. One such disease is Lowe disease (or oculocerebrorenal syndrome, because it affects the eyes, brain, and kidneys). In Lowe disease, there is a deficiency in an enzyme localized to the Golgi apparatus. Children with Lowe disease are born with cataracts, typically develop kidney disease after the

first year of life, and may have impaired mental abilities.

A mutation on the X chromosome causes Lowe disease. The X chromosome is one of the two human sex chromosomes, as these chromosomes determine a person's sex. Females possess two X chromosomes while males possess one X and one Y chromosome. In females, the genes on only one of the two X chromosomes are expressed. Females who carry the Lowe disease gene on one of their X chromosomes are carriers and do not show symptoms of the disease. However, males only have one X chromosome and the genes on this chromosome are always expressed. Therefore, males will always have Lowe disease if their X chromosome carries the Lowe disease gene.

Geneticists have identified the mutated gene's location, as well as many other mutation locations that cause genetic diseases. Through prenatal testing, a woman can find out if the fetus she is carrying may be afflicted with one of several genetic diseases.

Geneticists analyze prenatal genetic test results and may counsel pregnant women on available options. They may also conduct genetic research that leads to new drugs or foods, or perform DNA analyses for forensic investigations.

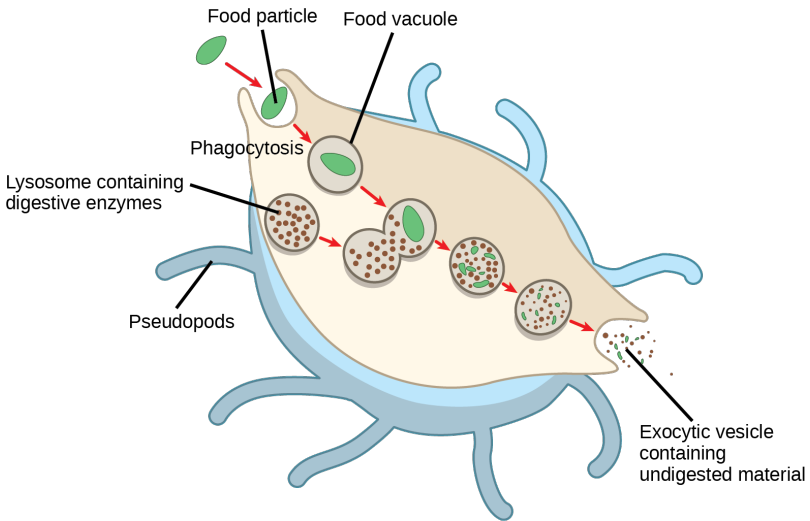
A macrophage has engulfed (phagocytized) a

potentially pathogenic bacterium and then fuses with lysosomes within the cell to destroy the pathogen. Other organelles are present in the cell but for simplicity we do not show them.

Lysosomes

In addition to their role as the digestive component and organelle-recycling facility of animal cells, lysosomes are part of the endomembrane system. Lysosomes also use their hydrolytic enzymes to destroy pathogens (disease-causing organisms) that might enter the cell. A good example of this occurs in macrophages, a group of white blood cells which are part of your body's immune system. In a process that scientists call phagocytosis or endocytosis, a section of the macrophage's plasma membrane invaginates (folds in) and engulfs a pathogen. The invaginated section, with the pathogen inside, then pinches itself off from the plasma membrane and becomes a vesicle. The vesicle fuses with a lysosome. The lysosome's hydrolytic enzymes then destroy the pathogen ([\[link\]](#)).

Phagocytosis



Section Summary

The endomembrane system includes the nuclear envelope, lysosomes, vesicles, the ER, and Golgi apparatus, as well as the plasma membrane. These cellular components work together to modify, package, tag, and transport proteins and lipids that form the membranes.

The RER modifies proteins and synthesizes phospholipids in cell membranes. The SER synthesizes carbohydrates, lipids, and steroid hormones; engages in the detoxification of medications and poisons; and stores calcium ions. Sorting, tagging, packaging, and distributing lipids and proteins take place in the Golgi apparatus.

Budding RER and Golgi membranes create lysosomes. Lysosomes digest macromolecules, recycle worn-out organelles, and destroy pathogens.

Visual Connection Questions

[\[link\]](#) If a peripheral membrane protein were synthesized in the lumen (inside) of the ER, would it end up on the inside or outside of the plasma membrane?

[\[link\]](#) It would end up on the outside. After the vesicle passes through the Golgi apparatus and fuses with the plasma membrane, it turns inside out.

Review Questions

Which of the following is not a component of the endomembrane system?

1. mitochondrion
2. Golgi apparatus
3. endoplasmic reticulum

4. lysosome

A

The process by which a cell engulfs a foreign particle is known as:

1. endosymbiosis
 2. phagocytosis
 3. hydrolysis
 4. membrane synthesis
-

B

Which of the following is most likely to have the greatest concentration of smooth endoplasmic reticulum?

1. a cell that secretes enzymes
 2. a cell that destroys pathogens
 3. a cell that makes steroid hormones
 4. a cell that engages in photosynthesis
-

C

Which of the following sequences correctly lists

in order the steps involved in the incorporation of a proteinaceous molecule within a cell?

1. protein synthesis of the protein on the ribosome; modification in the Golgi apparatus; packaging in the endoplasmic reticulum; tagging in the vesicle
2. synthesis of the protein on the lysosome; tagging in the Golgi; packaging in the vesicle; distribution in the endoplasmic reticulum
3. synthesis of the protein on the ribosome; modification in the endoplasmic reticulum; tagging in the Golgi; distribution via the vesicle
4. synthesis of the protein on the lysosome; packaging in the vesicle; distribution via the Golgi; tagging in the endoplasmic reticulum

C

Congenital disorders of glycosylation are a growing class of rare diseases. Which organelle would be most commonly involved in the glycoprotein disorder portion of the group?

1. RER
2. ribosomes
3. endosomes

4. Golgi apparatus

D

Critical Thinking Questions

In the context of cell biology, what do we mean by form follows function? What are at least two examples of this concept?

“Form follows function” refers to the idea that the function of a body part dictates the form of that body part. As an example, compare your arm to a bat’s wing. While the bones of the two correspond, the parts serve different functions in each organism and their forms have adapted to follow that function.

In your opinion, is the nuclear membrane part of the endomembrane system? Why or why not? Defend your answer.

Since the external surface of the nuclear membrane is continuous with the rough endoplasmic reticulum, which is part of the

endomembrane system, then it is correct to say that it is part of the system.

Glossary

endomembrane system

group of organelles and membranes in eukaryotic cells that work together modifying, packaging, and transporting lipids and proteins

endoplasmic reticulum (ER)

series of interconnected membranous structures within eukaryotic cells that collectively modify proteins and synthesize lipids

Golgi apparatus

eukaryotic organelle comprised of a series of stacked membranes that sorts, tags, and packages lipids and proteins for distribution

rough endoplasmic reticulum (RER)

region of the endoplasmic reticulum that is studded with ribosomes and engages in protein modification and phospholipid synthesis

smooth endoplasmic reticulum (SER)

region of the endoplasmic reticulum that has few or no ribosomes on its cytoplasmic

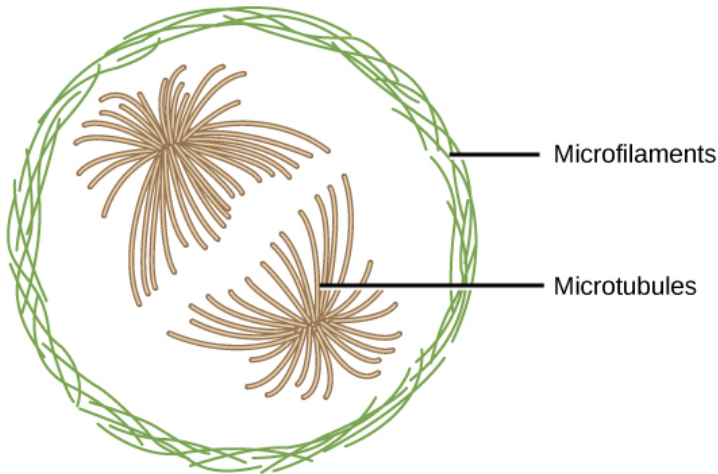
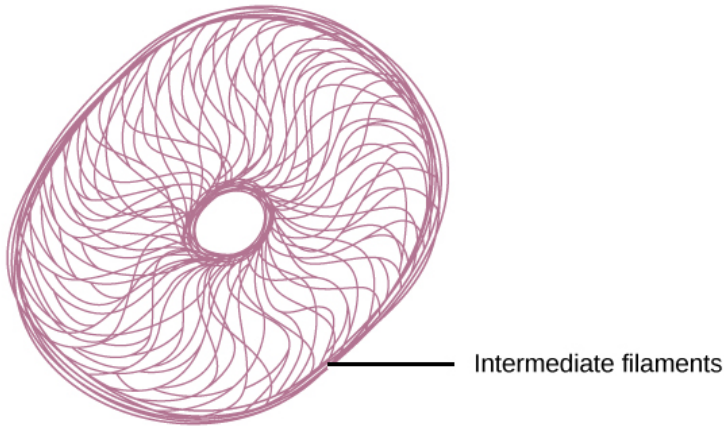
surface and synthesizes carbohydrates, lipids, and steroid hormones; detoxifies certain chemicals (like pesticides, preservatives, medications, and environmental pollutants), and stores calcium ions

The Cytoskeleton

By the end of this section, you will be able to do the following:

- Describe the cytoskeleton
- Compare the roles of microfilaments, intermediate filaments, and microtubules
- Compare and contrast cilia and flagella
- Summarize the differences among the components of prokaryotic cells, animal cells, and plant cells

If you were to remove all the organelles from a cell, would the plasma membrane and the cytoplasm be the only components left? No. Within the cytoplasm, there would still be ions and organic molecules, plus a network of protein fibers that help maintain the cell's shape, secure some organelles in specific positions, allow cytoplasm and vesicles to move within the cell, and enable cells within multicellular organisms to move. Collectively, scientists call this network of protein fibers the **cytoskeleton**. There are three types of fibers within the cytoskeleton: microfilaments, intermediate filaments, and microtubules ([\[link\]](#)). Here, we will examine each. Microfilaments thicken the cortex around the cell's inner edge. Like rubber bands, they resist tension. There are microtubules in the cell's interior where they maintain their shape by resisting compressive forces. There are intermediate filaments throughout the cell that hold organelles in place.

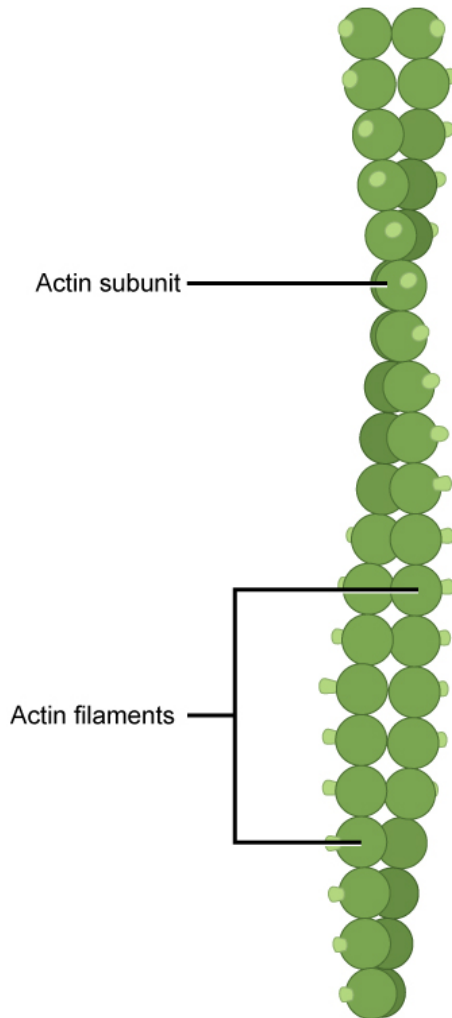


Two intertwined actin strands comprise microfilaments.

Microfilaments

Of the three types of protein fibers in the cytoskeleton, **microfilaments** are the narrowest. They function in cellular movement, have a diameter of about 7 nm, and are comprised of two globular protein intertwined strands, which we call

actin ([\[link\]](#)). For this reason, we also call microfilaments actin filaments.



ATP powers actin to assemble its filamentous form, which serves as a track for the movement of a motor protein we call myosin. This enables actin to engage in cellular events requiring motion, such as cell division in eukaryotic cells and cytoplasmic streaming, which is the cell cytoplasm's circular

movement in plant cells. Actin and myosin are plentiful in muscle cells. When your actin and myosin filaments slide past each other, your muscles contract.

Microfilaments also provide some rigidity and shape to the cell. They can depolymerize (disassemble) and reform quickly, thus enabling a cell to change its shape and move. White blood cells (your body's infection-fighting cells) make good use of this ability. They can move to an infection site and phagocytize the pathogen.

Link to Learning

To see an example of a white blood cell in action, watch a short time-lapse video of the cell capturing two bacteria. It engulfs one and then moves on to the other.

https://www.openstax.org/l/chasing_bacteria

Intermediate filaments consist of several intertwined strands of fibrous proteins.

Intermediate Filaments

Several strands of fibrous proteins that are wound together comprise intermediate filaments ([link]).

Cytoskeleton elements get their name from the fact that their diameter, 8 to 10 nm, is between those of microfilaments and microtubules.



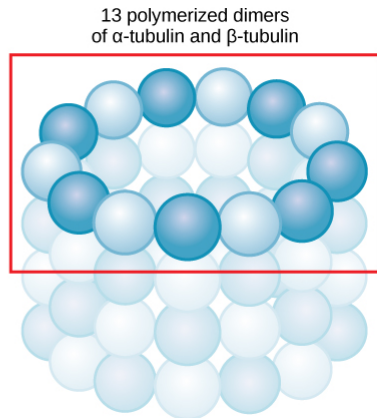
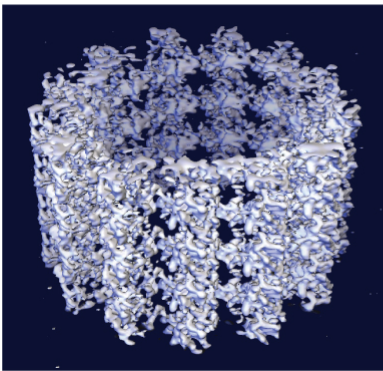
Intermediate filaments have no role in cell movement. Their function is purely structural. They bear tension, thus maintaining the cell's shape, and anchor the nucleus and other organelles in place. [\[link\]](#) shows how intermediate filaments create a supportive scaffolding inside the cell.

The intermediate filaments are the most diverse group of cytoskeletal elements. Several fibrous protein types are in the intermediate filaments. You are probably most familiar with keratin, the fibrous protein that strengthens your hair, nails, and the skin's epidermis.

Microtubules are hollow. Their walls consist of 13 polymerized dimers of α -tubulin and β -tubulin (right image). The left image shows the tube's molecular structure. This transmission electron micrograph of two flagella shows the microtubules' 9 + 2 array: nine microtubule doublets surround a single microtubule doublet. (credit: modification of work by Dartmouth Electron Microscope Facility, Dartmouth College; scale-bar data from Matt Russell)

Microtubules

As their name implies, microtubules are small hollow tubes. Polymerized dimers of α -tubulin and β -tubulin, two globular proteins, comprise the microtubule's walls ([\[link\]](#)). With a diameter of about 25 nm, **microtubules** are cytoskeletons' widest components. They help the cell resist compression, provide a track along which vesicles move through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. Like microfilaments, microtubules can disassemble and reform quickly.

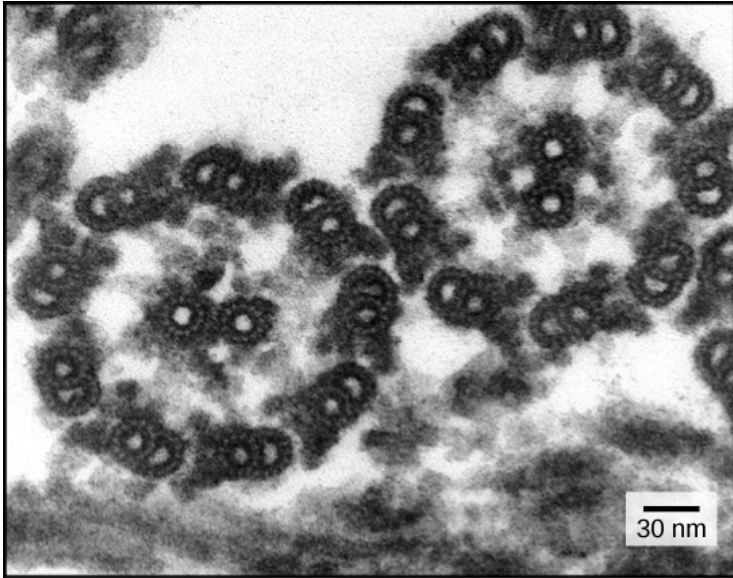


Microtubules are also the structural elements of flagella, cilia, and centrioles (the latter are the centrosome's two perpendicular bodies). In animal cells, the centrosome is the microtubule-organizing center. In eukaryotic cells, flagella and cilia are quite different structurally from their counterparts in prokaryotes, as we discuss below.

Flagella and Cilia

The **flagella** (singular = flagellum) are long, hair-like structures that extend from the plasma membrane and enable an entire cell to move (for example, sperm, *Euglena*, and some prokaryotes). When present, the cell has just one flagellum or a few flagella. However, when **cilia** (singular = cilium) are present, many of them extend along the plasma membrane's entire surface. They are short, hair-like structures that move entire cells (such as paramecia) or substances along the cell's outer surface (for example, the cilia of cells lining the Fallopian tubes that move the ovum toward the uterus, or cilia lining the cells of the respiratory tract that trap particulate matter and move it toward your nostrils.)

Despite their differences in length and number, flagella and cilia share a common structural arrangement of microtubules called a “9 + 2 array.” This is an appropriate name because a single flagellum or cilium is made of a ring of nine microtubule doublets, surrounding a single microtubule doublet in the center ([\[link\]](#)).



You have now completed a broad survey of prokaryotic and eukaryotic cell components. For a summary of cellular components in prokaryotic and eukaryotic cells, see [\[link\]](#).

Components of Prokaryotic and Eukaryotic Cells

Cell Component	Function	Present in	Present in	Present in Plant

	Prokaryotic Cells?	Eukaryotic Cells?	Animal Cells?	Plant Cells?
--	--------------------	-------------------	---------------	--------------

Plasma membrane	Separates cell from external environment; controls passage of organic molecules, ions, water, oxygen, and wastes into and out of cell	Yes	Yes	Yes
-----------------	---	-----	-----	-----

Cytoplasm	Provides turgor pressure to plant cells as fluid inside the central vacuole; site of many metabolic reactions; medium	Yes	Yes	Yes
-----------	---	-----	-----	-----

in which
organelles
are found

Nucleolus	Darkened area within the nucleus where ribosomal subunits are synthesized.	No	Yes	Yes
-----------	--	----	-----	-----

Darkened area within the nucleus where ribosomal subunits are synthesized.

Nucleus	Cell organelle that houses DNA and directs synthesis of ribosomes and proteins	No	Yes	Yes
---------	--	----	-----	-----

Cell organelle that houses DNA and directs synthesis of ribosomes and proteins

Ribosomes	Protein synthesis	Yes	Yes	Yes
-----------	-------------------	-----	-----	-----

Protein synthesis

Mitochondria	ATP production/cellular respiration	No	Yes	Yes
--------------	-------------------------------------	----	-----	-----

ATP production/cellular respiration

Peroxisomes	Oxidize and thus	No	Yes	Yes
-------------	------------------	----	-----	-----

Oxidize and thus

break
down
fatty acids
and
amino
acids, and
detoxify
poisons

Vesicles
and
vacuoles

Storage
and
transport;
digestive
function
in plant
cells

No

Yes

Yes

Centrosomes

Unspecified
role in
cell
division in
animal
cells;
microtubule
source in
animal
cells

No

Yes

No

Lysosomes

Digestion
of
macromolecules;
recycling
of worn-
out

No

Yes

Some

organelles

Cell wall	Protection, structural support, and maintenance of cell shape	Yes, primarily peptidoglycan	No	Yes, primarily cellulose
Chloroplast	Photosynthesis	Yes	No	Yes
Endoplasmic reticulum	Modifies and synthesizes proteins and lipids	No	Yes	Yes
Golgi apparatus	Modifies, sorts, tags, packages, and distributes lipids and proteins	No	Yes	Yes
Cytoskeleton	Maintains cell's shape, secures organelles in specific positions, allows cytoplasm	Yes	Yes	Yes

and vesicles to move within cell, and enables unicellular organisms to move independently

Flagella

Cellular locomotion

Some

Some

No, except for some plant sperm cells
No

Cilia

Cellular locomotion, movement of particles along plasma membrane's extracellular surface, and filtration

Some

Some

Section Summary

The cytoskeleton has three different protein element types. From narrowest to widest, they are the microfilaments (actin filaments), intermediate filaments, and microtubules. Biologists often associate microfilaments with myosin. They provide rigidity and shape to the cell and facilitate cellular movements. Intermediate filaments bear tension and anchor the nucleus and other organelles in place. Microtubules help the cell resist compression, serve as tracks for motor proteins that move vesicles through the cell, and pull replicated chromosomes to opposite ends of a dividing cell. They are also the structural element of centrioles, flagella, and cilia.

Review Questions

Which of the following have the ability to disassemble and reform quickly?

1. microfilaments and intermediate filaments
2. microfilaments and microtubules
3. intermediate filaments and microtubules
4. only intermediate filaments

Which of the following do not play a role in intracellular movement?

1. microfilaments and intermediate filaments
 2. microfilaments and microtubules
 3. intermediate filaments and microtubules
 4. only intermediate filaments
-

D

In humans, ____ are used to move a cell within its environment while ____ are used to move the environment relative to the cell.

1. cilia, pseudopodia
 2. flagella; cilia
 3. microtubules; flagella
 4. microfilaments; microtubules
-

B

Critical Thinking Questions

What are the similarities and differences between the structures of centrioles and

flagella?

Centrioles and flagella are alike in that they are made up of microtubules. In centrioles, two rings of nine microtubule “triplets” are arranged at right angles to one another. This arrangement does not occur in flagella.

How do cilia and flagella differ?

Cilia and flagella are alike in that they are made up of microtubules. Cilia are short, hair-like structures that exist in large numbers and usually cover the entire surface of the plasma membrane. Flagella, in contrast, are long, hair-like structures; when flagella are present, a cell has just one or two.

Describe how microfilaments and microtubules are involved in the phagocytosis and destruction of a pathogen by a macrophage.

A macrophage engulfs a pathogen by rearranging its actin microfilaments to bend the plasma membrane around the pathogen. Once the pathogen is sealed in an endosome inside the macrophage, the vesicle is walked along

microtubules until it combines with a lysosome to digest the pathogen.

Compare and contrast the boundaries that plant, animal, and bacteria cells use to separate themselves from their surrounding environment.

All three cell types have a plasma membrane that borders the cytoplasm on its interior side. In animal cells, the exterior side of the plasma membrane is in contact with the extracellular environment. However, in plant and bacteria cells, a cell wall surrounds the outside of the plasma membrane. In plants, the cell wall is made of cellulose, while in bacteria the cell wall is made of peptidoglycan. Gram-negative bacteria also have an additional capsule made of lipopolysaccharides that surrounds their cell wall.

Glossary

cilium

(plural = cilia) short, hair-like structure that extends from the plasma membrane in large numbers and functions to move an entire cell or move substances along the cell's outer surface

cytoskeleton

protein fiber network that collectively maintains the cell's shape, secures some organelles in specific positions, allows cytoplasm and vesicles to move within the cell, and enables unicellular organisms to move independently

flagellum

(plural = flagella) long, hair-like structure that extends from the plasma membrane and moves the cell

intermediate filament

cytoskeletal component, comprised of several fibrous protein intertwined strands, that bears tension, supports cell-cell junctions, and anchors cells to extracellular structures

microfilament

the cytoskeleton system's narrowest element; it provides rigidity and shape to the cell and enables cellular movements

microtubule

the cytoskeleton system's widest element; it helps the cell resist compression, provides a track along which vesicles move through the cell, pulls replicated chromosomes to opposite ends of a dividing cell, and is the structural element of centrioles, flagella, and cilia

Connections between Cells and Cellular Activities

By the end of this section, you will be able to do the following:

- Describe the extracellular matrix
- List examples of the ways that plant cells and animal cells communicate with adjacent cells
- Summarize the roles of tight junctions, desmosomes, gap junctions, and plasmodesmata

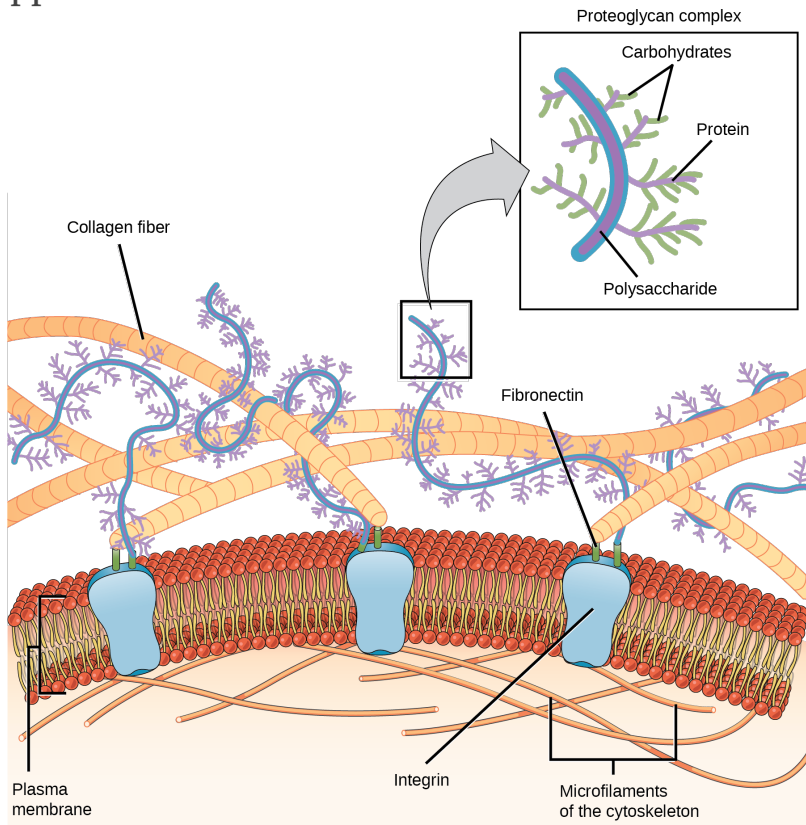
You already know that tissue is a group of similar cells working together. As you might expect, if cells are to work together, they must communicate with each other, just as you need to communicate with others if you work on a group project. Let's take a look at how cells communicate with each other.

The extracellular matrix consists of a network of proteins and carbohydrates.

Extracellular Matrix of Animal Cells

While cells in most multicellular organisms release materials into the extracellular space, animal cells will be discussed as an example. The primary components of these materials are proteins, and the most abundant protein is collagen. Collagen fibers are interwoven with proteoglycans, which are carbohydrate-containing protein molecules. Collectively, we call these materials the

extracellular matrix ([\[link\]](#)). Not only does the extracellular matrix hold the cells together to form a tissue, but it also allows the cells within the tissue to communicate with each other. How can this happen?



Cells have protein receptors on their plasma membranes' extracellular surfaces. When a molecule within the matrix binds to the receptor, it changes the receptor's molecular structure. The receptor, in turn, changes the microfilaments' conformation positioned just inside the plasma membrane. These conformational changes induce chemical signals

inside the cell that reach the nucleus and turn “on” or “off” the transcription of specific DNA sections, which affects the associated protein production, thus changing the activities within the cell.

Blood clotting provides an example of the extracellular matrix's role in cell communication. When the cells lining a blood vessel are damaged, they display a protein receptor, which we call tissue factor. When tissue factor binds with another factor in the extracellular matrix, it causes platelets to adhere to the damaged blood vessel's wall, stimulates the adjacent smooth muscle cells in the blood vessel to contract (thus constricting the blood vessel), and initiates a series of steps that stimulate the platelets to produce clotting factors.

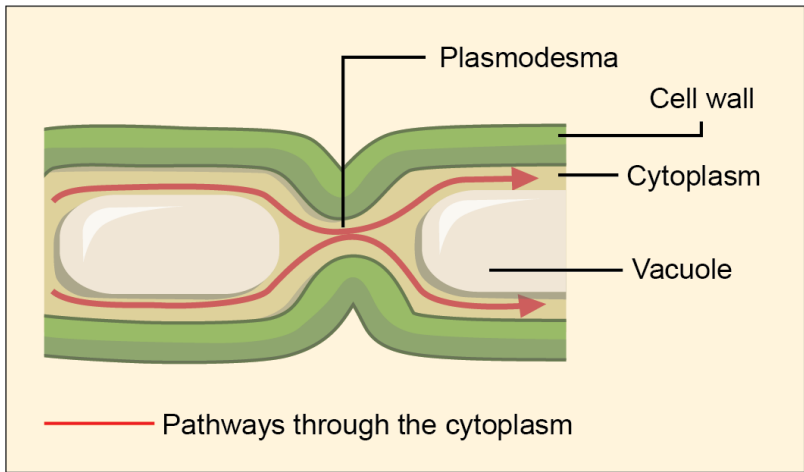
A plasmodesma is a channel between two adjacent plant cells' cell walls. Plasmodesmata allow materials to pass from one plant cell's cytoplasm to an adjacent cell's cytoplasm. Tight junctions form watertight connections between adjacent animal cells. Proteins create tight junction adherence. (credit: modification of work by Mariana Ruiz Villareal) A desmosome forms a very strong spot weld between cells. Linking cadherins and intermediate filaments create it. (credit: modification of work by Mariana Ruiz Villareal) A gap junction is a protein-lined pore that allows water and small molecules to pass between adjacent animal cells. (credit: modification of work by Mariana Ruiz Villareal)

Intercellular Junctions

Cells can also communicate with each other via direct contact, or intercellular junctions. There are differences in the ways that plant and animal and fungal cells communicate. Plasmodesmata are junctions between plant cells; whereas, animal cell contacts include tight junctions, gap junctions, and desmosomes.

Plasmodesmata

In general, long stretches of the plasma membranes of neighboring plant cells cannot touch one another because the cell wall that surrounds each cell separates them ([\[link\]](#)). How then, can a plant transfer water and other soil nutrients from its roots, through its stems, and to its leaves? Such transport uses the vascular tissues (xylem and phloem) primarily. There also exist structural modifications, which we call **plasmodesmata** (singular = plasmodesma). Numerous channels that pass between adjacent plant cells' cell walls connect their cytoplasm, and enable transport of materials from cell to cell, and thus throughout the plant ([\[link\]](#)).



Tight Junctions

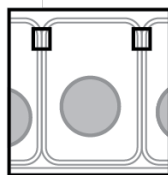
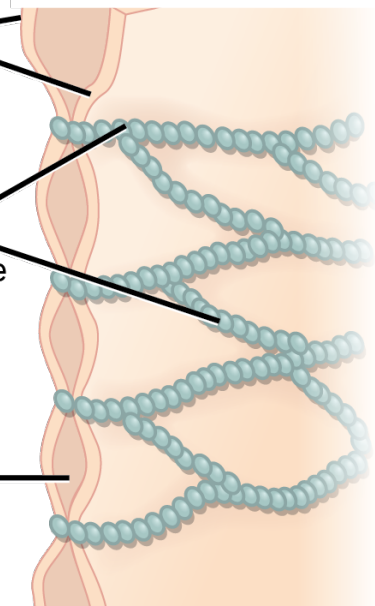
A **tight junction** is a watertight seal between two adjacent animal cells ([\[link\]](#)). Proteins (predominantly two proteins called claudins and occludins) tightly hold the cells against each other.

Tight junction

Adjacent
plasma
membranes

Strands of
transmembrane
proteins

Intercellular
space



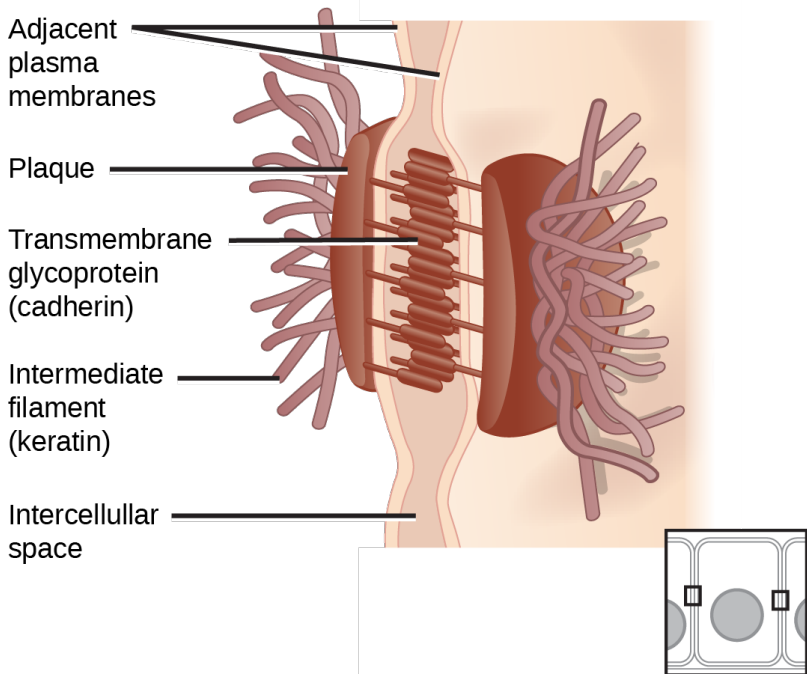
This tight adherence prevents materials from leaking between the cells; tight junctions are typically found in epithelial tissues that line internal organs and cavities, and comprise most of the skin. For example, the tight junctions of the epithelial cells lining your urinary bladder prevent urine from leaking out into the extracellular space.

Desmosomes

Also only in animal cells are **desmosomes**, which act like spot welds between adjacent epithelial cells ([\[link\]](#)). Cadherins, short proteins in the plasma membrane connect to intermediate filaments to

create desmosomes. The cadherins connect two adjacent cells and maintain the cells in a sheet-like formation in organs and tissues that stretch, like the skin, heart, and muscles.

Desmosome



Gap Junctions

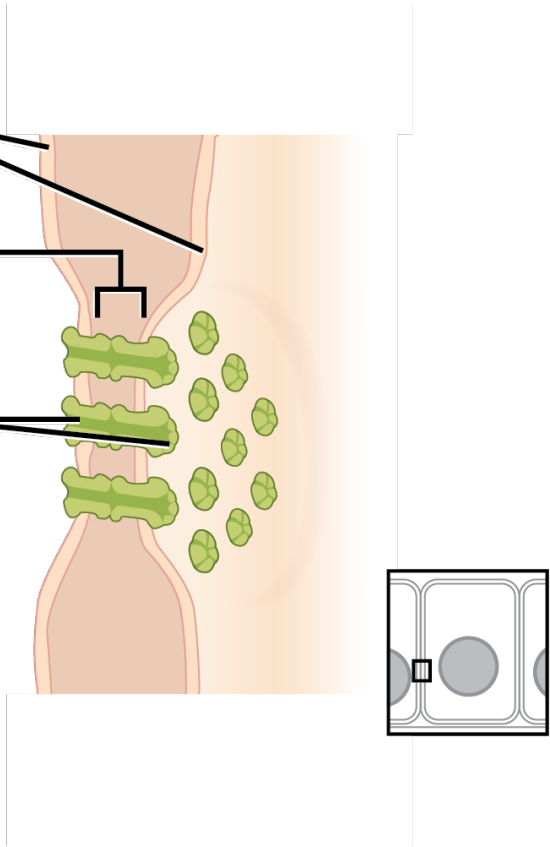
Gap junctions in animal cells are like plasmodesmata in plant cells in that they are channels between adjacent cells that allow for transporting ions, nutrients, and other substances that enable cells to communicate ([\[link\]](#)). Structurally, however, gap junctions and plasmodesmata differ.

Gap junction

Adjacent
plasma
membranes

Gap between
cells

Connexons
(composed of
connexins)



Gap junctions develop when a set of six proteins (connexins) in the plasma membrane arrange themselves in an elongated donut-like configuration - a connexon. When the connexon's pores ("doughnut holes") in adjacent animal cells align, a channel between the two cells forms. Gap junctions are particularly important in cardiac muscle. The electrical signal for the muscle to contract passes efficiently through gap junctions, allowing the heart muscle cells to contract in tandem.

Link to Learning

To conduct a virtual microscopy lab and review the parts of a cell, work through the steps of this [interactive assignment](#).

Section Summary

Animal cells communicate via their extracellular matrices and are connected to each other via tight junctions, desmosomes, and gap junctions. Plant cells are connected and communicate with each other via plasmodesmata.

When protein receptors on the plasma membrane's surface of an animal cell bind to a substance in the extracellular matrix, a chain of reactions begins that changes activities taking place within the cell. Plasmodesmata are channels between adjacent plant cells, while gap junctions are channels between adjacent animal cells. However, their structures are quite different. A tight junction is a watertight seal between two adjacent cells, while a desmosome acts like a spot weld.

Review Questions

Which of the following are only in plant cells?

1. gap junctions
 2. desmosomes
 3. plasmodesmata
 4. tight junctions
-

C

The key components of desmosomes are cadherins and _____.

1. actin
 2. microfilaments
 3. intermediate filaments
 4. microtubules
-

C

Diseased animal cells may produce molecules that activate death cascades to kill the cells in a controlled manner. Why would neighboring healthy cells also die?

1. The death molecule is passed through desmosomes.
2. The death molecule is passed through

plasmodesmata.

3. The death molecule disrupts the extracellular matrix.
4. The death molecule passes through gap junctions.

D

Critical Thinking Questions

How does the structure of a plasmodesma differ from that of a gap junction?

They differ because plant cell walls are rigid. Plasmodesmata, which a plant cell needs for transportation and communication, are able to allow movement of really large molecules. Gap junctions are necessary in animal cells for transportation and communication.

Explain how the extracellular matrix functions.

The extracellular matrix functions in support and attachment for animal tissues. It also functions in the healing and growth of the

tissue.

Pathogenic *E. coli* have recently been shown to degrade tight junction proteins during infection. How would this provide an advantage to the bacteria?

E. coli infections generally cause food poisoning, meaning that the invading bacteria cross from the lumen of the gut into the rest of the body. Tight junctions hold the epithelial layer that lines the digestive tract together so that the material that crosses into the body is tightly regulated. One way *E. coli* can avoid this regulation is to destroy the tight junctions so that it can enter the body between the epithelial cells, rather than having to go through the cells.

Glossary

desmosome

linkages between adjacent epithelial cells that form when cadherins in the plasma membrane attach to intermediate filaments

extracellular matrix

material secreted from animal or fungal cells that provides mechanical protection and

anchoring for the cells in the tissue

gap junction

channel between two adjacent animal cells that allows ions, nutrients, and low molecular weight substances to pass between cells, enabling the cells to communicate

plasmodesma

(plural = plasmodesmata) channel that passes between adjacent plant cells' cell walls, connects their cytoplasm, and allows transporting of materials from cell to cell

tight junction

protein adherence that creates a firm seal between two adjacent animal cells

Introduction

class = "introduction" Despite its seeming hustle and bustle, Grand Central Station functions with a high level of organization: People and objects move from one location to another, they cross or are contained within certain boundaries, and they provide a constant flow as part of larger activity. Analogously, a plasma membrane's functions involve movement within the cell and across boundaries' activities. (credit: modification of work by Randy Le'Moine)



The plasma membrane, the cell membrane, has many functions, but the most basic one is to define the cell's borders and keep the cell functional. The plasma membrane is selectively permeable. This means that the membrane allows some materials to freely enter or leave the cell, while other materials cannot move freely, but require a specialized structure, and occasionally, even energy investment for crossing.

Components and Structure

By the end of this section, you will be able to do the following:

- Understand the cell membrane fluid mosaic model
- Describe phospholipid, protein, and carbohydrate functions in membranes
- Discuss membrane fluidity

A cell's plasma membrane defines the cell, outlines its borders, and determines the nature of its interaction with its environment (see [\[link\]](#) for a summary). Cells exclude some substances, take in others, and excrete still others, all in controlled quantities. The plasma membrane must be very flexible to allow certain cells, such as red and white blood cells, to change shape as they pass through narrow capillaries. These are the more obvious plasma membrane functions. In addition, the plasma membrane's surface carries markers that allow cells to recognize one another, which is vital for tissue and organ formation during early development, and which later plays a role in the immune response's “self” versus “non-self” distinction.

Among the most sophisticated plasma membrane functions is the ability for complex, integral proteins, receptors to transmit signals. These proteins act both as extracellular input receivers and as intracellular processing activators. These

membrane receptors provide extracellular attachment sites for effectors like hormones and growth factors, and they activate intracellular response cascades when their effectors are bound. Occasionally, viruses hijack receptors (HIV, human immunodeficiency virus, is one example) that use them to gain entry into cells, and at times, the genes encoding receptors become mutated, causing the signal transduction process to malfunction with disastrous consequences.

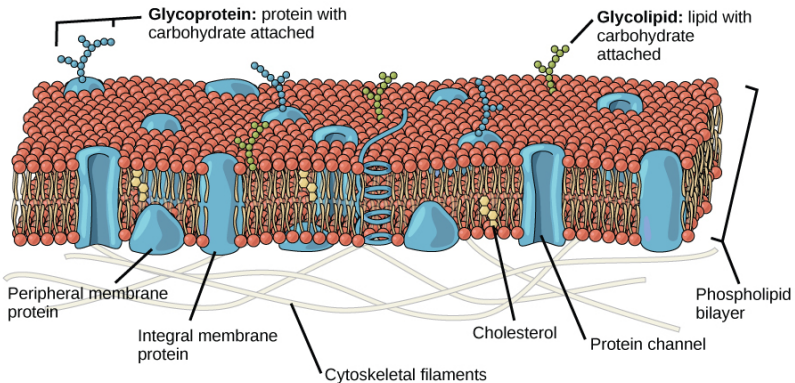
The plasma membrane fluid mosaic model describes the plasma membrane as a fluid combination of phospholipids, cholesterol, and proteins. Carbohydrates attached to lipids (glycolipids) and to proteins (glycoproteins) extend from the membrane's outward-facing surface. A hydrophilic head and two hydrophobic tails comprise this phospholipid molecule. The hydrophilic head group consists of a phosphate-containing group attached to a glycerol molecule. The hydrophobic tails, each containing either a saturated or an unsaturated fatty acid, are long hydrocarbon chains. In an aqueous solution, phospholipids usually arrange themselves with their polar heads facing outward and their hydrophobic tails facing inward. (credit: modification of work by Mariana Ruiz Villareal) Integral membrane proteins may have one or more alpha-helices that span the membrane (examples 1 and 2), or they may have beta-sheets that span the membrane (example 3). (credit: "Foobar"/Wikimedia Commons)

Fluid Mosaic Model

Scientists identified the plasma membrane in the 1890s, and its chemical components in 1915. The principal components they identified were lipids and proteins. In 1935, Hugh Davson and James Danielli proposed the plasma membrane's structure. This was the first model that others in the scientific community widely accepted. It was based on the plasma membrane's “railroad track” appearance in early electron micrographs. Davson and Danielli theorized that the plasma membrane's structure resembles a sandwich. They made the analogy of proteins to bread, and lipids to the filling. In the 1950s, advances in microscopy, notably transmission electron microscopy (TEM), allowed researchers to see that the plasma membrane's core consisted of a double, rather than a single, layer. In 1972, S.J. Singer and Garth L. Nicolson proposed a new model that provides microscopic observations and better explains plasma membrane function.

The explanation, the **fluid mosaic model**, has evolved somewhat over time, but it still best accounts for plasma membrane structure and function as we now understand them. The fluid mosaic model describes the plasma membrane structure as a mosaic of components—including phospholipids, cholesterol, proteins, and carbohydrates—that gives the membrane a fluid character. Plasma membranes range from 5 to 10

nm in thickness. For comparison, human red blood cells, visible via light microscopy, are approximately 8 μm wide, or approximately 1,000 times wider than a plasma membrane. The membrane does look a bit like a sandwich ([link]).



A plasma membrane's principal components are lipids (phospholipids and cholesterol), proteins, and carbohydrates attached to some of the lipids and proteins. A phospholipid is a molecule consisting of glycerol, two fatty acids, and a phosphate-linked head group. Cholesterol, another lipid comprised of four fused carbon rings, is situated alongside the phospholipids in the membrane's core. The protein, lipid, and carbohydrate proportions in the plasma membrane vary with cell type, but for a typical human cell, protein accounts for about 50 percent of the composition by mass, lipids (of all types) account for about 40 percent, and carbohydrates comprise the remaining 10 percent. However, protein and lipid concentration varies with different cell membranes. For example, myelin, an outgrowth of specialized cells' membrane that insulates the

peripheral nerves' axons, contains only 18 percent protein and 76 percent lipid. The mitochondrial inner membrane contains 76 percent protein and only 24 percent lipid. The plasma membrane of human red blood cells is 30 percent lipid.

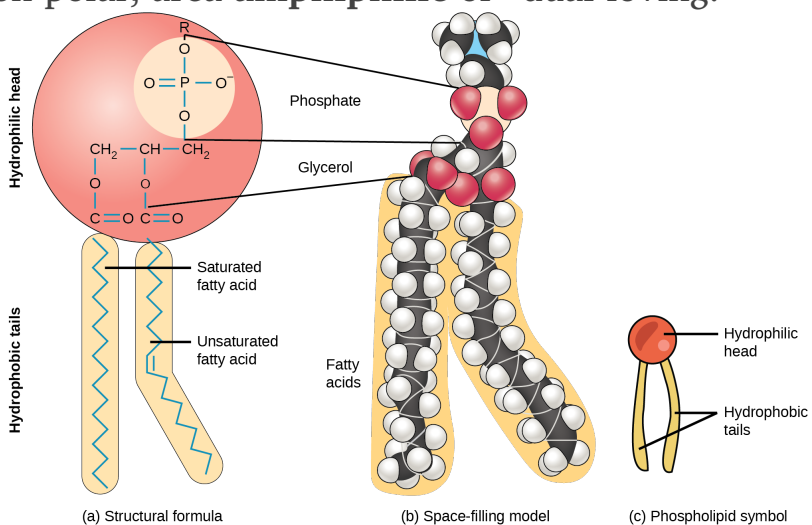
Carbohydrates are present only on the plasma membrane's exterior surface and are attached to proteins, forming **glycoproteins**, or attached to lipids, forming **glycolipids**.

Phospholipids

The membrane's main fabric comprises amphiphilic, phospholipid molecules. The **hydrophilic** or “water-loving” areas of these molecules (which look like a collection of balls in an artist’s rendition of the model) ([\[link\]](#)) are in contact with the aqueous fluid both inside and outside the cell. **Hydrophobic**, or water-hating molecules, tend to be non-polar. They interact with other non-polar molecules in chemical reactions, but generally do not interact with polar molecules. When placed in water, hydrophobic molecules tend to form a ball or cluster. The phospholipids' hydrophilic regions form hydrogen bonds with water and other polar molecules on both the cell's exterior and interior. Thus, the membrane surfaces that face the cell's interior and exterior are hydrophilic. In contrast, the cell membrane's interior is hydrophobic and will not interact with water. Therefore, phospholipids form an excellent two-layer cell membrane that separates fluid within

the cell from the fluid outside the cell.

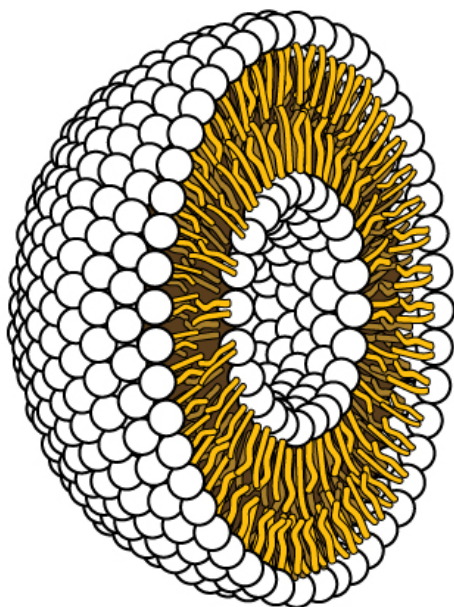
A phospholipid molecule ([\[link\]](#)) consists of a three-carbon glycerol backbone with two fatty acid molecules attached to carbons 1 and 2, and a phosphate-containing group attached to the third carbon. This arrangement gives the overall molecule a head area (the phosphate-containing group), which has a polar character or negative charge, and a tail area (the fatty acids), which has no charge. The head can form hydrogen bonds, but the tail cannot. Scientists call a molecule with a positively or negatively charged area and an uncharged, or non-polar, area **amphiphilic** or “dual-loving.”



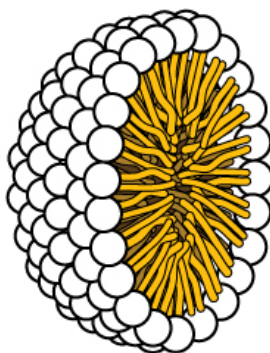
This characteristic is vital to the plasma membrane's structure because, in water, phospholipids arrange themselves with their hydrophobic tails facing each other and their hydrophilic heads facing out. In this way, they form a lipid bilayer—a double layered

phospholipid barrier that separates the water and other materials on one side from the water and other materials on the other side. Phospholipids heated in an aqueous solution usually spontaneously form small spheres or droplets (micelles or liposomes), with their hydrophilic heads forming the exterior and their hydrophobic tails on the inside ([link]).

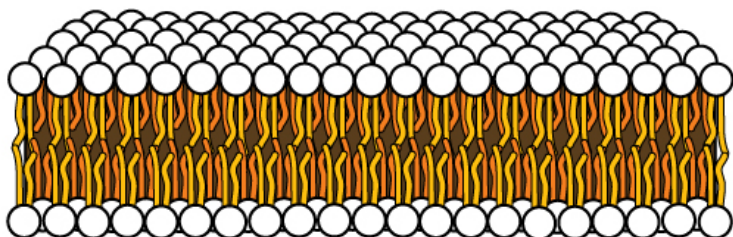
Lipid-bilayer sphere



Single-layer lipid sphere

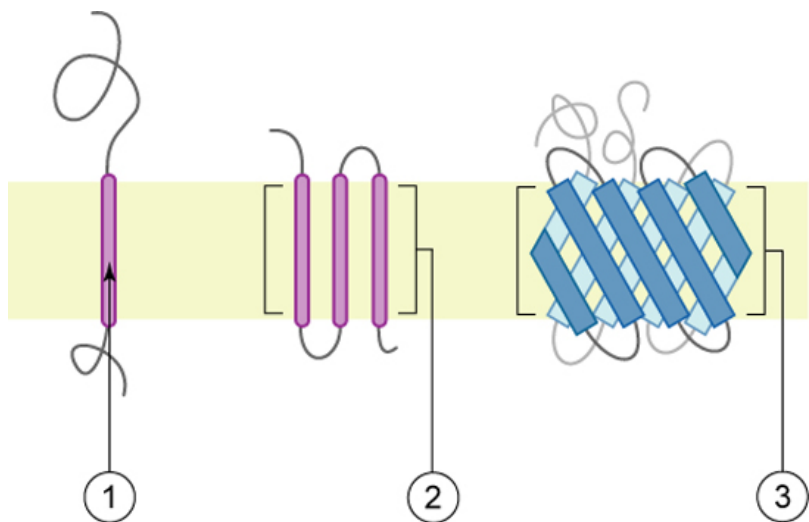


Lipid-bilayer sheet



Proteins

Proteins comprise the plasma membranes' second major component. **Integral proteins**, or integrins, as their name suggests, integrate completely into the membrane structure, and their hydrophobic membrane-spanning regions interact with the phospholipid bilayer's hydrophobic region ([\[link\]](#)). Single-pass integral membrane proteins usually have a hydrophobic transmembrane segment that consists of 20–25 amino acids. Some span only part of the membrane—associating with a single layer—while others stretch from one side to the other, and are exposed on either side. Up to 12 single protein segments comprise some complex proteins, which are extensively folded and embedded in the membrane ([\[link\]](#)). This protein type has a hydrophilic region or regions, and one or several mildly hydrophobic regions. This arrangement of protein regions orients the protein alongside the phospholipids, with the protein's hydrophobic region adjacent to the phospholipids' tails and the protein's hydrophilic region or regions protruding from the membrane and in contact with the cytosol or extracellular fluid.



Peripheral proteins are on the membranes' exterior and interior surfaces, attached either to integral proteins or to phospholipids. Peripheral proteins, along with integral proteins, may serve as enzymes, as structural attachments for the cytoskeleton's fibers, or as part of the cell's recognition sites. Scientists sometimes refer to these as “cell-specific” proteins. The body recognizes its own proteins and attacks foreign proteins associated with invasive pathogens.

Carbohydrates

Carbohydrates are the third major plasma membrane component. They are always on the cells' exterior surface and are bound either to proteins (forming glycoproteins) or to lipids (forming glycolipids) ([link](#)). These carbohydrate chains may consist of 2–60 monosaccharide units and can be

either straight or branched. Along with peripheral proteins, carbohydrates form specialized sites on the cell surface that allow cells to recognize each other. These sites have unique patterns that allow for cell recognition, much the way that the facial features unique to each person allow individuals to recognize him or her. This recognition function is very important to cells, as it allows the immune system to differentiate between body cells (“self”) and foreign cells or tissues (“non-self”). Similar glycoprotein and glycolipid types are on the surfaces of viruses and may change frequently, preventing immune cells from recognizing and attacking them.

We collectively refer to these carbohydrates on the cell's exterior surface—the carbohydrate components of both glycoproteins and glycolipids—as the glycocalyx (meaning “sugar coating”). The glycocalyx is highly hydrophilic and attracts large amounts of water to the cell's surface. This aids in the cell's interaction with its watery environment and in the cell's ability to obtain substances dissolved in the water. As we discussed above, the glycocalyx is also important for cell identification, self/non-self determination, and embryonic development, and is used in cell to cell attachments to form tissues.

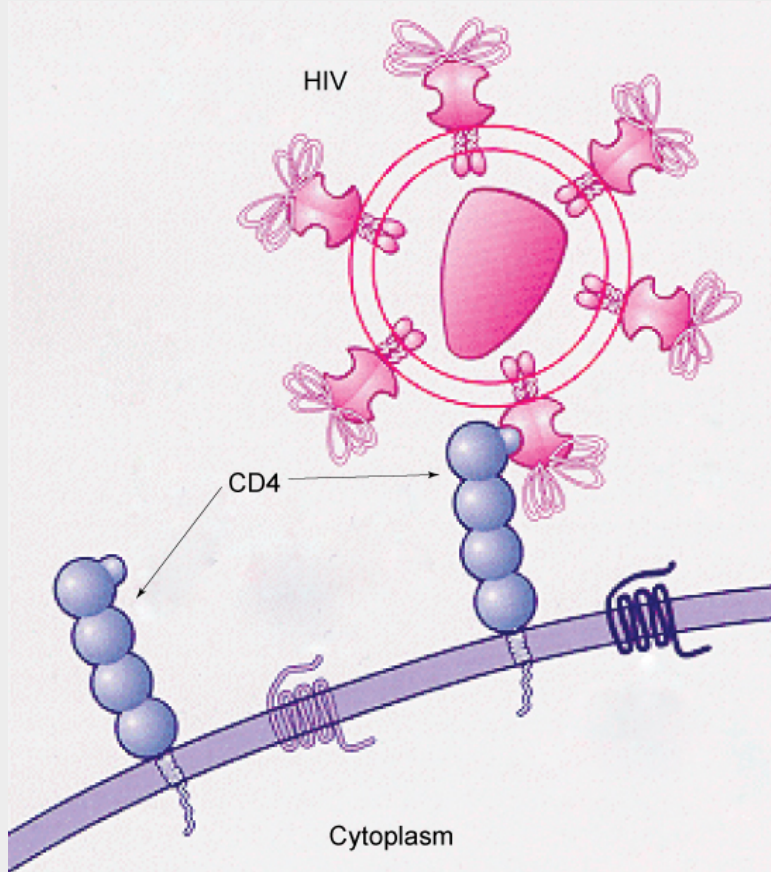
Evolution Connection

How Viruses Infect Specific Organs

Glycoprotein and glycolipid patterns on the cells' surfaces give many viruses an opportunity for infection. HIV and hepatitis viruses infect only specific organs or cells in the human body. HIV is able to penetrate the plasma membranes of a subtype of lymphocytes called T-helper cells, as well as some monocytes and central nervous system cells. The hepatitis virus attacks liver cells. These viruses are able to invade these cells, because the cells have binding sites on their surfaces that are specific to and compatible with certain viruses ([\[link\]](#)). Other recognition sites on the virus's surface interact with the human immune system, prompting the body to produce antibodies. Antibodies are made in response to the antigens or proteins associated with invasive pathogens, or in response to foreign cells, such as might occur with an organ transplant. These same sites serve as places for antibodies to attach and either destroy or inhibit the virus' activity. Unfortunately, these recognition sites on HIV change at a rapid rate because of mutations, making an effective vaccine against the virus very difficult, as the virus evolves and adapts. A person infected with HIV will quickly develop different populations, or variants, of the virus that differences in these recognition sites distinguish. This rapid change of surface markers decreases the effectiveness of the person's immune system in attacking the virus, because the antibodies will not recognize the surface patterns'

new variations. In the case of HIV, the problem is compounded because the virus specifically infects and destroys cells involved in the immune response, further incapacitating the host.

HIV binds to the CD4 receptor, a glycoprotein on T cell surfaces. (credit: modification of work by NIH, NIAID)



Membrane Fluidity

The membrane's mosaic characteristic helps to illustrate its nature. The integral proteins and lipids exist in the membrane as separate but loosely attached molecules. These resemble the separate, multicolored tiles of a mosaic picture, and they float, moving somewhat with respect to one another. The membrane is not like a balloon, however, that can expand and contract; rather, it is fairly rigid and can burst if penetrated or if a cell takes in too much water. However, because of its mosaic nature, a very fine needle can easily penetrate a plasma membrane without causing it to burst, and the membrane will flow and self-seal when one extracts the needle.

The membrane's mosaic characteristics explain some but not all of its fluidity. There are two other factors that help maintain this fluid characteristic. One factor is the nature of the phospholipids themselves. In their saturated form, the fatty acids in phospholipid tails are saturated with bound hydrogen atoms. There are no double bonds between adjacent carbon atoms. This results in tails that are relatively straight. In contrast, unsaturated fatty acids do not contain a maximal number of hydrogen atoms, but they do contain some double bonds between adjacent carbon atoms. A double bond results in a bend in the carbon string of approximately 30 degrees ([\[link\]](#)).

Thus, if decreasing temperatures compress saturated fatty acids with their straight tails, they press in on each other, making a dense and fairly rigid membrane. If unsaturated fatty acids are compressed, the “kinks” in their tails elbow adjacent phospholipid molecules away, maintaining some space between the phospholipid molecules. This “elbow room” helps to maintain fluidity in the membrane at temperatures at which membranes with saturated fatty acid tails in their phospholipids would “freeze” or solidify. The membrane's relative fluidity is particularly important in a cold environment. A cold environment usually compresses membranes comprised largely of saturated fatty acids, making them less fluid and more susceptible to rupturing. Many organisms (fish are one example) are capable of adapting to cold environments by changing the proportion of unsaturated fatty acids in their membranes in response to lower temperature.

Link to Learning

Visit this [site](#) to see animations of the membranes' fluidity and mosaic quality.

Animals have an additional membrane constituent that assists in maintaining fluidity. Cholesterol,

which lies alongside the phospholipids in the membrane, tends to dampen temperature effects on the membrane. Thus, this lipid functions as a buffer, preventing lower temperatures from inhibiting fluidity and preventing increased temperatures from increasing fluidity too much. Thus, cholesterol extends, in both directions, the temperature range in which the membrane is appropriately fluid and consequently functional. Cholesterol also serves other functions, such as organizing clusters of transmembrane proteins into lipid rafts.

Plasma Membrane Components and Functions	
Component	Location
Phospholipid	Main membrane fabric
Cholesterol	Attached between phospholipids and between the two phospholipid layers
Integral proteins (for example, integrins)	Embedded within the phospholipid layer(s); may or may not penetrate through both layers
Peripheral proteins	On the phospholipid bilayer's inner or outer

Carbohydrates
(components of
glycoproteins and
glycolipids)

surface; not embedded
within the phospholipids
Generally attached to
proteins on the outside
membrane layer

Career Connection

Immunologist

The variations in peripheral proteins and carbohydrates that affect a cell's recognition sites are of prime interest in immunology. In developing vaccines, researchers have been able to conquer many infectious diseases, such as smallpox, polio, diphtheria, and tetanus.

Immunologists are the physicians and scientists who research and develop vaccines, as well as treat and study allergies or other immune problems.

Some immunologists study and treat autoimmune problems (diseases in which a person's immune system attacks his or her own cells or tissues, such as lupus) and immunodeficiencies, whether acquired (such as acquired immunodeficiency syndrome, or AIDS) or hereditary (such as severe combined immunodeficiency, or SCID).

Immunologists also help treat organ transplantation patients, who must have their immune systems suppressed so that their bodies will not reject a transplanted organ. Some

immunologists work to understand natural immunity and the effects of a person's environment on it. Others work on questions about how the immune system affects diseases such as cancer. In the past, researchers did not understand the importance of having a healthy immune system in preventing cancer.

To work as an immunologist, one must have a PhD or MD. In addition, immunologists undertake at least two to three years of training in an accredited program and must pass the American Board of Allergy and Immunology exam. Immunologists must possess knowledge of the human body's function as they relate to issues beyond immunization, and knowledge of pharmacology and medical technology, such as medications, therapies, test materials, and surgical procedures.

Section Summary

Modern scientists refer to the plasma membrane as the fluid mosaic model. A phospholipid bilayer comprises the plasma membrane, with hydrophobic, fatty acid tails in contact with each other. The membrane's landscape is studded with proteins, some which span the membrane. Some of these proteins serve to transport materials into or out of

the cell. Carbohydrates are attached to some of the proteins and lipids on the membrane's outward-facing surface, forming complexes that function to identify the cell to other cells. The membrane's fluid nature is due to temperature, fatty acid tail configuration (some kinked by double bonds), cholesterol presence embedded in the membrane, and the mosaic nature of the proteins and protein-carbohydrate combinations, which are not firmly fixed in place. Plasma membranes enclose and define the cells' borders. Not static, they are dynamic and constantly in flux.

Review Questions

Which plasma membrane component can be either found on its surface or embedded in the membrane structure?

1. protein
2. cholesterol
3. carbohydrate
4. phospholipid

A

Which characteristic of a phospholipid

contributes to the fluidity of the membrane?

1. its head
 2. cholesterol
 3. a saturated fatty acid tail
 4. double bonds in the fatty acid tail
-

D

What is the primary function of carbohydrates attached to the exterior of cell membranes?

1. identification of the cell
 2. flexibility of the membrane
 3. strengthening the membrane
 4. channels through membrane
-

A

A scientist compares the plasma membrane composition of an animal from the Mediterranean coast with one from the Mojave Desert. Which hypothesis is most likely to be correct?

1. The cells from the Mediterranean coast animal will have more fluid plasma membranes.

2. The cells from the Mojave Desert animal will have a higher cholesterol concentration in the plasma membranes.
 3. The cells' plasma membranes will be indistinguishable.
 4. The cells from the Mediterranean coast animal will have a higher glycoprotein content, while the cells from the Mojave Desert animal will have a higher lipoprotein content.
-

B

Critical Thinking Questions

Why is it advantageous for the cell membrane to be fluid in nature?

The fluid characteristic of the cell membrane allows greater flexibility to the cell than it would if the membrane were rigid. It also allows the motion of membrane components, required for some types of membrane transport.

Why do phospholipids tend to spontaneously

orient themselves into something resembling a membrane?

The hydrophobic, nonpolar regions must align with each other in order for the structure to have minimal potential energy and, consequently, higher stability. The fatty acid tails of the phospholipids cannot mix with water, but the phosphate “head” of the molecule can. Thus, the head orients to water, and the tail to other lipids.

How can a cell use an extracellular peripheral protein as the receptor to transmit a signal into the cell?

Peripheral proteins can bind to other molecules in the extracellular space. However, they cannot directly transmit a signal to the inside of the cell since they do not have a transmembrane domain (region that goes through the plasma membrane to the inside of the cell). They must associate with integral membrane proteins in order to pass the signal to the inside of the cell.

Glossary

amphiphilic

molecule possessing a polar or charged area and a nonpolar or uncharged area capable of interacting with both hydrophilic and hydrophobic environments

fluid mosaic model

describes the plasma membrane's structure as a mosaic of components including phospholipids, cholesterol, proteins, glycoproteins, and glycolipids (sugar chains attached to proteins or lipids, respectively), resulting in a fluid character (fluidity)

glycolipid

combination of carbohydrates and lipids

glycoprotein

combination of carbohydrates and proteins

hydrophilic

molecule with the ability to bond with water; "water-loving"

hydrophobic

molecule that does not have the ability to bond with water; "water-hating"

integral protein

protein integrated into the membrane structure that interacts extensively with the membrane lipids' hydrocarbon chains and

often spans the membrane

peripheral protein

protein at the plasma membrane's surface
either on its exterior or interior side

Passive Transport

By the end of this section, you will be able to do the following:

- Explain why and how passive transport occurs
- Understand the osmosis and diffusion processes
- Define tonicity and its relevance to passive transport

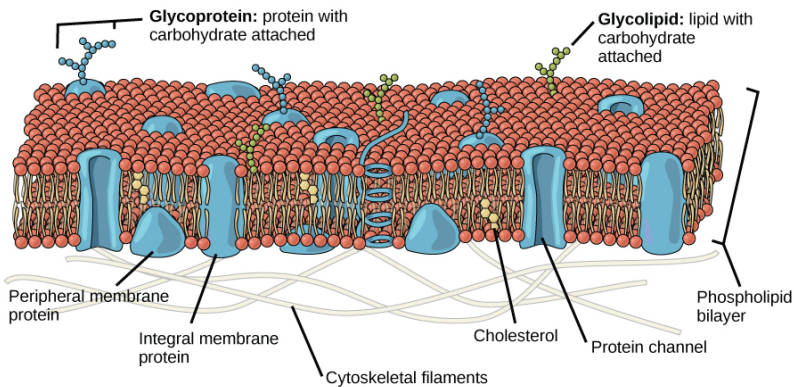
Plasma membranes must allow certain substances to enter and leave a cell, and prevent some harmful materials from entering and some essential materials from leaving. In other words, plasma membranes are **selectively permeable**—they allow some substances to pass through, but not others. If they were to lose this selectivity, the cell would no longer be able to sustain itself, and it would be destroyed. Some cells require larger amounts of specific substances. They must have a way of obtaining these materials from extracellular fluids. This may happen passively, as certain materials move back and forth, or the cell may have special mechanisms that facilitate transport. Some materials are so important to a cell that it spends some of its energy, hydrolyzing adenosine triphosphate (ATP), to obtain these materials. Red blood cells use some of their energy doing just that. Most cells spend the majority of their energy to maintain an imbalance of sodium and potassium ions between the cell's interior and exterior, as well as on protein synthesis.

The most direct forms of membrane transport are passive. **Passive transport** is a naturally occurring phenomenon and does not require the cell to exert any of its energy to accomplish the movement. In passive transport, substances move from an area of higher concentration to an area of lower concentration. A physical space in which there is a single substance concentration range has a **concentration gradient**.

The plasma membrane's exterior surface is not identical to its interior surface.

Selective Permeability

Plasma membranes are asymmetric: the membrane's interior is not identical to its exterior. There is a considerable difference between the array of phospholipids and proteins between the two leaflets that form a membrane. On the membrane's interior, some proteins serve to anchor the membrane to cytoskeleton's fibers. There are peripheral proteins on the membrane's exterior that bind extracellular matrix elements. Carbohydrates, attached to lipids or proteins, are also on the plasma membrane's exterior surface. These carbohydrate complexes help the cell bind required substances in the extracellular fluid. This adds considerably to plasma membrane's selective nature ([\[link\]](#)).



Recall that plasma membranes are amphiphilic: They have hydrophilic and hydrophobic regions. This characteristic helps move some materials through the membrane and hinders the movement of others. Non-polar and lipid-soluble material with a low molecular weight can easily slip through the membrane's hydrophobic lipid core. Substances such as the fat-soluble vitamins A, D, E, and K readily pass through the plasma membranes in the digestive tract and other tissues. Fat-soluble drugs and hormones also gain easy entry into cells and readily transport themselves into the body's tissues and organs. Oxygen and carbon dioxide molecules have no charge and pass through membranes by simple diffusion.

Polar substances present problems for the membrane. While some polar molecules connect easily with the cell's outside, they cannot readily pass through the plasma membrane's lipid core. Additionally, while small ions could easily slip through the spaces in the membrane's mosaic, their

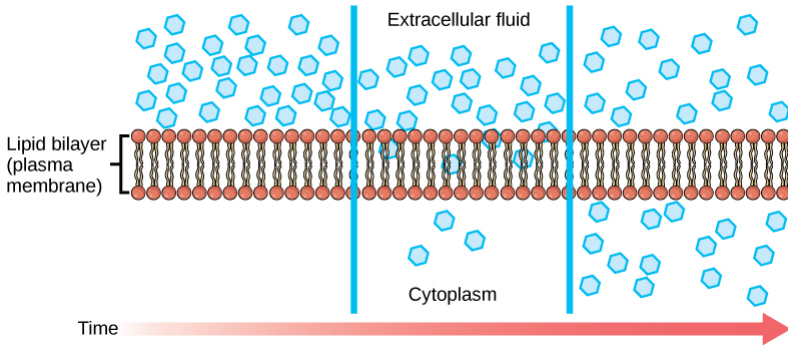
charge prevents them from doing so. Ions such as sodium, potassium, calcium, and chloride must have special means of penetrating plasma membranes. Simple sugars and amino acids also need the help of various transmembrane proteins (channels) to transport themselves across plasma membranes.

Diffusion through a permeable membrane moves a substance from a high concentration area (extracellular fluid, in this case) down its concentration gradient (into the cytoplasm). (credit: modification of work by Mariana Ruiz Villareal)

Diffusion

Diffusion is a passive process of transport. A single substance moves from a high concentration to a low concentration area until the concentration is equal across a space. You are familiar with diffusion of substances through the air. For example, think about someone opening a bottle of ammonia in a room filled with people. The ammonia gas is at its highest concentration in the bottle. Its lowest concentration is at the room's edges. The ammonia vapor will diffuse, or spread away, from the bottle, and gradually, increasingly more people will smell the ammonia as it spreads. Materials move within the cell's cytosol by diffusion, and certain materials move through the plasma membrane by diffusion ([\[link\]](#)). Diffusion expends no energy. On the contrary, concentration gradients are a form of potential energy, which dissipates as the gradient is

eliminated.



Each separate substance in a medium, such as the extracellular fluid, has its own concentration gradient, independent of other materials' concentration gradients. In addition, each substance will diffuse according to that gradient. Within a system, there will be different diffusion rates of various substances in the medium.

Factors That Affect Diffusion

Molecules move constantly in a random manner, at a rate that depends on their mass, their environment, and the amount of thermal energy they possess, which in turn is a function of temperature. This movement accounts for molecule diffusion through whatever medium in which they are localized. A substance moves into any space available to it until it evenly distributes itself throughout. After a substance has diffused completely through a space, removing its concentration gradient, molecules will still move

around in the space, but there will be no *net* movement of the number of molecules from one area to another. We call this lack of a concentration gradient in which the substance has no net movement dynamic equilibrium. While diffusion will go forward in the presence of a substance's concentration gradient, several factors affect the diffusion rate.

- Extent of the concentration gradient: The greater the difference in concentration, the more rapid the diffusion. The closer the distribution of the material gets to equilibrium, the slower the diffusion rate.
- Mass of the molecules diffusing: Heavier molecules move more slowly; therefore, they diffuse more slowly. The reverse is true for lighter molecules.
- Temperature: Higher temperatures increase the energy and therefore the molecules' movement, increasing the diffusion rate. Lower temperatures decrease the molecules' energy, thus decreasing the diffusion rate.
- Solvent density: As the density of a solvent increases, the diffusion rate decreases. The molecules slow down because they have a more difficult time passing through the denser medium. If the medium is less dense, diffusion increases. Because cells primarily use diffusion to move materials within the cytoplasm, any increase in the cytoplasm's density will inhibit

the movement of the materials. An example of this is a person experiencing dehydration. As the body's cells lose water, the diffusion rate decreases in the cytoplasm, and the cells' functions deteriorate. Neurons tend to be very sensitive to this effect. Dehydration frequently leads to unconsciousness and possibly coma because of the decrease in diffusion rate within the cells.

- Solubility: As we discussed earlier, nonpolar or lipid-soluble materials pass through plasma membranes more easily than polar materials, allowing a faster diffusion rate.
- Surface area and plasma membrane thickness: Increased surface area increases the diffusion rate; whereas, a thicker membrane reduces it.
- Distance travelled: The greater the distance that a substance must travel, the slower the diffusion rate. This places an upper limitation on cell size. A large, spherical cell will die because nutrients or waste cannot reach or leave the cell's center, respectively. Therefore, cells must either be small in size, as in the case of many prokaryotes, or be flattened, as with many single-celled eukaryotes.

A variation of diffusion is the process of filtration. In filtration, material moves according to its concentration gradient through a membrane. Sometimes pressure enhances the diffusion rate, causing the substances to filter more rapidly. This

occurs in the kidney, where blood pressure forces large amounts of water and accompanying dissolved substances, or **solutes**, out of the blood and into the renal tubules. The diffusion rate in this instance is almost totally dependent on pressure. One of the effects of high blood pressure is the appearance of protein in the urine, which abnormally high pressure "squeezes through".

Facilitated transport moves substances down their concentration gradients. They may cross the plasma membrane with the aid of channel proteins. (credit: modification of work by Mariana Ruiz Villareal)
Some substances are able to move down their concentration gradient across the plasma membrane with the aid of carrier proteins. Carrier proteins change shape as they move molecules across the membrane. (credit: modification of work by Mariana Ruiz Villareal)

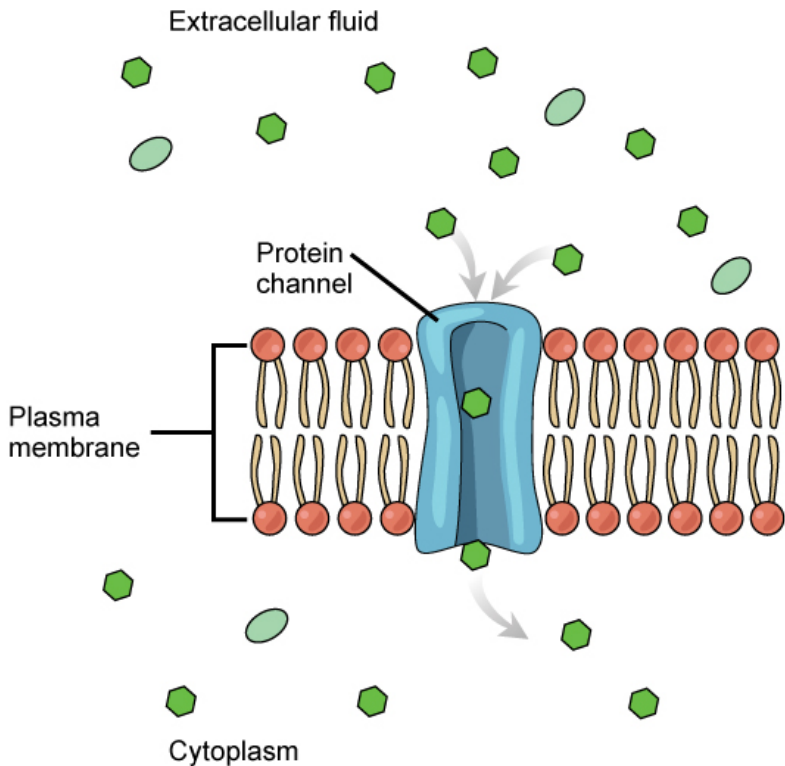
Facilitated transport

In **facilitated transport**, or facilitated diffusion, materials diffuse across the plasma membrane with the help of membrane proteins. A concentration gradient exists that would allow these materials to diffuse into the cell without expending cellular energy. However, these materials are polar molecule ions that the cell membrane's hydrophobic parts repel. Facilitated transport proteins shield these materials from the membrane's repulsive force, allowing them to diffuse into the cell.

The transported material first attaches to protein or glycoprotein receptors on the plasma membrane's exterior surface. This allows removal of material from the extracellular fluid that the cell needs. The substances then pass to specific integral proteins that facilitate their passage. Some of these integral proteins are collections of beta-pleated sheets that form a pore or channel through the phospholipid bilayer. Others are carrier proteins which bind with the substance and aid its diffusion through the membrane.

Channels

The integral proteins involved in facilitated transport are **transport proteins**, and they function as either channels for the material or carriers. In both cases, they are transmembrane proteins. Channels are specific for the transported substance. **Channel proteins** have hydrophilic domains exposed to the intracellular and extracellular fluids. In addition, they have a hydrophilic channel through their core that provides a hydrated opening through the membrane layers ([\[link\]](#)). Passage through the channel allows polar compounds to avoid the plasma membrane's nonpolar central layer that would otherwise slow or prevent their entry into the cell. **Aquaporins** are channel proteins that allow water to pass through the membrane at a very high rate.

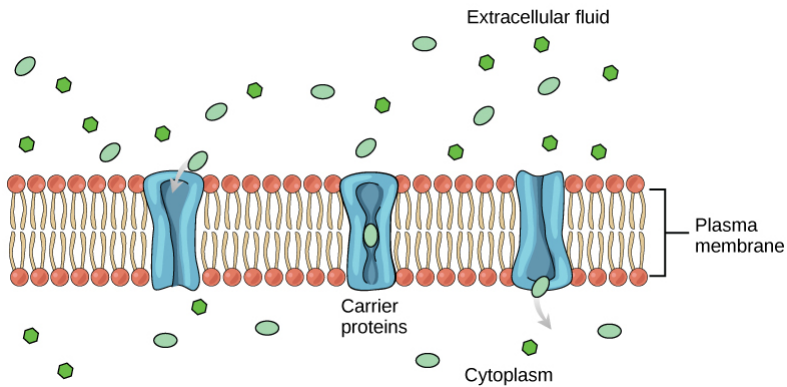


Channel proteins are either open at all times or they are “gated,” which controls the channel's opening. When a particular ion attaches to the channel protein it may control the opening, or other mechanisms or substances may be involved. In some tissues, sodium and chloride ions pass freely through open channels; whereas, in other tissues a gate must open to allow passage. An example of this occurs in the kidney, where there are both channel forms in different parts of the renal tubules. Cells involved in transmitting electrical impulses, such as nerve and muscle cells, have gated channels for sodium, potassium, and calcium in their

membranes. Opening and closing these channels changes the relative concentrations on opposing sides of the membrane of these ions, resulting in facilitating electrical transmission along membranes (in the case of nerve cells) or in muscle contraction (in the case of muscle cells).

Carrier Proteins

Another type of protein embedded in the plasma membrane is a **carrier protein**. This aptly named protein binds a substance and, thus triggers a change of its own shape, moving the bound molecule from the cell's outside to its interior ([\[link\]](#)). Depending on the gradient, the material may move in the opposite direction. Carrier proteins are typically specific for a single substance. This selectivity adds to the plasma membrane's overall selectivity. Scientists poorly understand the exact mechanism for the change of shape. Proteins can change shape when their hydrogen bonds are affected, but this may not fully explain this mechanism. Each carrier protein is specific to one substance, and there are a finite number of these proteins in any membrane. This can cause problems in transporting enough material for the cell to function properly. When all of the proteins are bound to their ligands, they are saturated and the rate of transport is at its maximum. Increasing the concentration gradient at this point will not result in an increased transport rate.



An example of this process occurs in the kidney. In one part, the kidney filters glucose, water, salts, ions, and amino acids that the body requires. This filtrate, which includes glucose, then reabsorbs in another part of the kidney. Because there are only a finite number of carrier proteins for glucose, if more glucose is present than the proteins can handle, the excess is not transported and the body excretes this through urine. In a diabetic individual, the term is “spilling glucose into the urine.” A different group of carrier proteins, glucose transport proteins, or GLUTs, are involved in transporting glucose and other hexose sugars through plasma membranes within the body.

Channel and carrier proteins transport material at different rates. Channel proteins transport much more quickly than carrier proteins. Channel proteins facilitate diffusion at a rate of tens of millions of molecules per second; whereas, carrier proteins work at a rate of a thousand to a million molecules per second.

In osmosis, water always moves from an area of higher water concentration to one of lower concentration. In the diagram, the solute cannot pass through the selectively permeable membrane, but the water can.

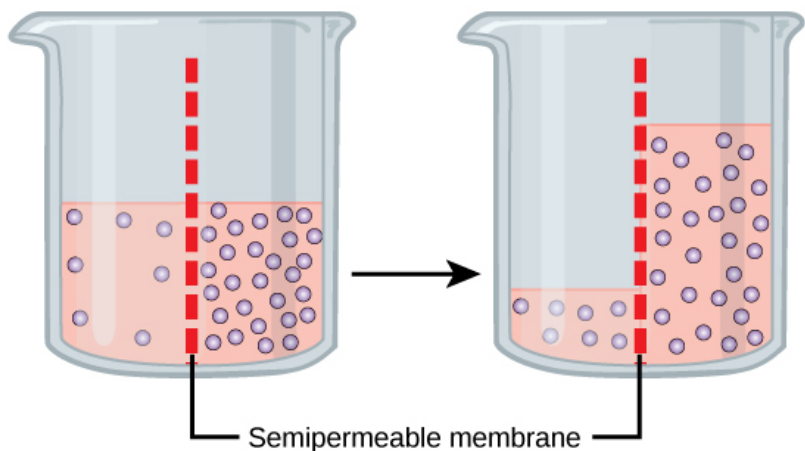
Osmosis

Osmosis is the movement of water through a semipermeable membrane according to the water's concentration gradient across the membrane, which is inversely proportional to the solutes' concentration. While diffusion transports material across membranes and within cells, osmosis transports *only water* across a membrane and the membrane limits the solutes' diffusion in the water. Not surprisingly, the aquaporins that facilitate water movement play a large role in osmosis, most prominently in red blood cells and the membranes of kidney tubules.

Mechanism

Osmosis is a special case of diffusion. Water, like other substances, moves from an area of high concentration to one of low concentration. An obvious question is what makes water move at all? Imagine a beaker with a semipermeable membrane separating the two sides or halves ([\[link\]](#)). On both sides of the membrane the water level is the same,

but there are different dissolved substance concentrations, or **solute**, that cannot cross the membrane (otherwise the solute crossing the membrane would balance concentrations on each side). If the solution's volume on both sides of the membrane is the same, but the solute's concentrations are different, then there are different amounts of water, the solvent, on either side of the membrane.



To illustrate this, imagine two full water glasses. One has a single teaspoon of sugar in it; whereas, the second one contains one-quarter cup of sugar. If the total volume of the solutions in both cups is the same, which cup contains more water? Because the large sugar amount in the second cup takes up much more space than the teaspoon of sugar in the first cup, the first cup has more water in it.

Returning to the beaker example, recall that it has a solute mixture on either side of the membrane. A

principle of diffusion is that the molecules move around and will spread evenly throughout the medium if they can. However, only the material capable of getting through the membrane will diffuse through it. In this example, the solute cannot diffuse through the membrane, but the water can. Water has a concentration gradient in this system. Thus, water will diffuse down its concentration gradient, crossing the membrane to the side where it is less concentrated. This diffusion of water through the membrane—osmosis—will continue until the water's concentration gradient goes to zero or until the water's hydrostatic pressure balances the osmotic pressure. Osmosis proceeds constantly in living systems.

Tonicity

Tonicity describes how an extracellular solution can change a cell's volume by affecting osmosis. A solution's tonicity often directly correlates with the solution's osmolarity. **Osmolarity** describes the solution's total solute concentration. A solution with low osmolarity has a greater number of water molecules relative to the number of solute particles. A solution with high osmolarity has fewer water molecules with respect to solute particles. In a situation in which a membrane permeable to water, though not to the solute separates two different osmolarities, water will move from the membrane's

side with lower osmolarity (and more water) to the side with higher osmolarity (and less water). This effect makes sense if you remember that the solute cannot move across the membrane, and thus the only component in the system that can move—the water—moves along its own concentration gradient. An important distinction that concerns living systems is that osmolarity measures the number of particles (which may be molecules) in a solution. Therefore, a solution that is cloudy with cells may have a lower osmolarity than a solution that is clear, if the second solution contains more dissolved molecules than there are cells.

Hypotonic Solutions

Scientists use three terms—hypotonic, isotonic, and hypertonic—to relate the cell's osmolarity to the extracellular fluid's osmolarity that contains the cells. In a **hypotonic** situation, the extracellular fluid has lower osmolarity than the fluid inside the cell, and water enters the cell. (In living systems, the point of reference is always the cytoplasm, so the prefix *hypo-* means that the extracellular fluid has a lower solute concentration, or a lower osmolarity, than the cell cytoplasm.) It also means that the extracellular fluid has a higher water concentration in the solution than does the cell. In this situation, water will follow its concentration gradient and enter the cell.

Hypertonic Solutions

As for a **hypertonic** solution, the prefix *hyper-* refers to the extracellular fluid having a higher osmolarity than the cell's cytoplasm; therefore, the fluid contains less water than the cell does. Because the cell has a relatively higher water concentration, water will leave the cell.

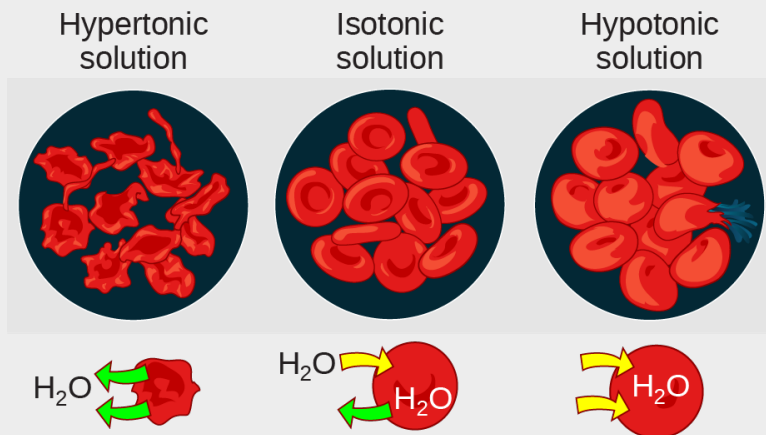
Isotonic Solutions

In an **isotonic** solution, the extracellular fluid has the same osmolarity as the cell. If the cell's osmolarity matches that of the extracellular fluid, there will be no net movement of water into or out of the cell, although water will still move in and out. Blood cells and plant cells in hypertonic, isotonic, and hypotonic solutions take on characteristic appearances ([\[link\]](#)).

Visual Connection

Osmotic pressure changes red blood cells' shape in hypertonic, isotonic, and hypotonic solutions.

(credit: Mariana Ruiz Villareal)



A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

Link to Learning

For a video illustrating the diffusion process in solutions, visit this [site](#).

The turgor pressure within a plant cell depends on the solution's tonicity in which it is bathed. (credit: modification of work by Mariana Ruiz Villareal) Without adequate water, the plant on the left has lost turgor pressure, visible in its wilting. Watering the plant (right) will restore the turgor pressure. (credit: Victor M. Vicente Selvas) A paramecium's

contractile vacuole, here visualized using bright field light microscopy at 480x magnification, continuously pumps water out of the organism's body to keep it from bursting in a hypotonic medium. (credit: modification of work by NIH; scale-bar data from Matt Russell)

Tonicity in Living Systems

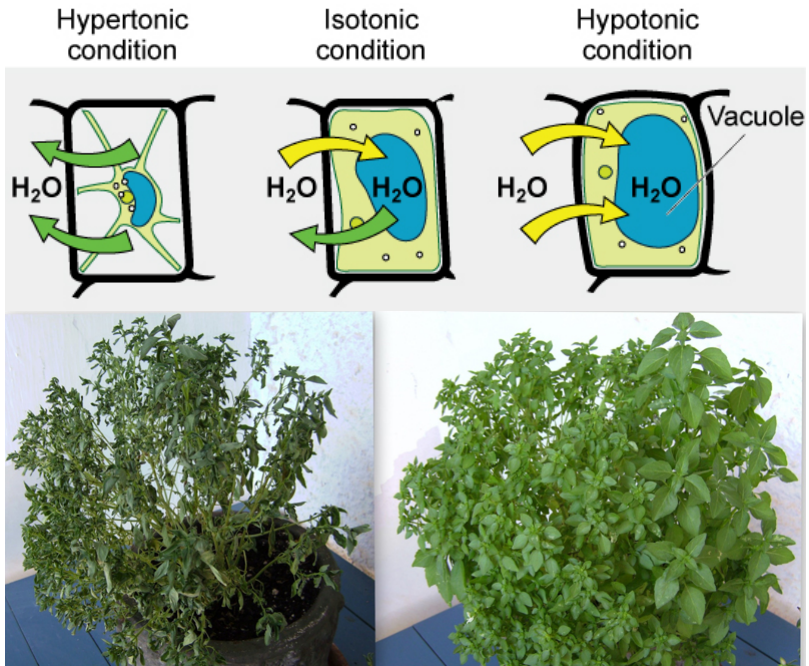
In a hypotonic environment, water enters a cell, and the cell swells. In an isotonic condition, the relative solute and solvent concentrations are equal on both membrane sides. There is no net water movement; therefore, there is no change in the cell's size. In a hypertonic solution, water leaves a cell and the cell shrinks. If either the hypo- or hyper- condition goes to excess, the cell's functions become compromised, and the cell may be destroyed.

A red blood cell will burst, or lyse, when it swells beyond the plasma membrane's capability to expand. Remember, the membrane resembles a mosaic, with discrete spaces between the molecules comprising it. If the cell swells, and the spaces between the lipids and proteins become too large, the cell will break apart.

In contrast, when excessive water amounts leave a red blood cell, the cell shrinks, or crenates. This has the effect of concentrating the solutes left in the cell, making the cytosol denser and interfering with

diffusion within the cell. The cell's ability to function will be compromised and may also result in the cell's death.

Various living things have ways of controlling the effects of osmosis—a mechanism we call osmoregulation. Some organisms, such as plants, fungi, bacteria, and some protists, have cell walls that surround the plasma membrane and prevent cell lysis in a hypotonic solution. The plasma membrane can only expand to the cell wall's limit, so the cell will not lyse. The cytoplasm in plants is always slightly hypertonic to the cellular environment, and water will always enter a cell if water is available. This water inflow produces turgor pressure, which stiffens the plant's cell walls ([link]). In nonwoody plants, turgor pressure supports the plant. Conversely, if you do not water the plant, the extracellular fluid will become hypertonic, causing water to leave the cell. In this condition, the cell does not shrink because the cell wall is not flexible. However, the cell membrane detaches from the wall and constricts the cytoplasm. We call this **plasmolysis**. Plants lose turgor pressure in this condition and wilt ([link]).



Tonicity is a concern for all living things. For example, paramecia and amoebas, which are protists that lack cell walls, have contractile vacuoles. This vesicle collects excess water from the cell and pumps it out, keeping the cell from lysing as it takes on water from its environment ([\[link\]](#)).



Many marine invertebrates have internal salt levels matched to their environments, making them isotonic with the water in which they live. Fish, however, must spend approximately five percent of their metabolic energy maintaining osmotic homeostasis. Freshwater fish live in an environment that is hypotonic to their cells. These fish actively take in salt through their gills and excrete diluted urine to rid themselves of excess water. Saltwater fish live in the reverse environment, which is hypertonic to their cells, and they secrete salt through their gills and excrete highly concentrated urine.

In vertebrates, the kidneys regulate the water amount in the body. Osmoreceptors are specialized cells in the brain that monitor solute concentration in the blood. If the solute levels increase beyond a certain range, a hormone releases that slows water loss through the kidney and dilutes the blood to safer levels. Animals also have high albumin concentrations, which the liver produces, in their blood. This protein is too large to pass easily through plasma membranes and is a major factor in controlling the osmotic pressures applied to tissues.

Section Summary

The passive transport forms, diffusion and osmosis, move materials of small molecular weight across

membranes. Substances diffuse from high to lower concentration areas, and this process continues until the substance evenly distributes itself in a system. In solutions containing more than one substance, each molecule type diffuses according to its own concentration gradient, independent of other substances diffusing. Many factors can affect the diffusion rate, such as concentration gradient, diffusing, particle sizes, and the system's temperature.

In living systems, the plasma membrane mediates substances diffusing in and out of cells. Some materials diffuse readily through the membrane, but others are hindered and only can pass through due to specialized proteins such as channels and transporters. The chemistry of living things occurs in aqueous solutions, and balancing the concentrations of those solutions is an ongoing problem. In living systems, diffusing some substances would be slow or difficult without membrane proteins that facilitate transport.

Visual Connection Questions

[\[link\]](#) A doctor injects a patient with what the doctor thinks is an isotonic saline solution. The patient dies, and an autopsy reveals that many

red blood cells have been destroyed. Do you think the solution the doctor injected was really isotonic?

[\[link\]](#) No, it must have been hypotonic as a hypotonic solution would cause water to enter the cells, thereby making them burst.

Review Questions

Water moves via osmosis _____.

1. throughout the cytoplasm
2. from an area with a high concentration of other solutes to a lower one
3. from an area with a high concentration of water to one of lower concentration
4. from an area with a low concentration of water to higher concentration

C

The principal force driving movement in diffusion is the _____.

1. temperature
 2. particle size
 3. concentration gradient
 4. membrane surface area
-

C

What problem is faced by organisms that live in fresh water?

1. Their bodies tend to take in too much water.
 2. They have no way of controlling their tonicity.
 3. Only salt water poses problems for animals that live in it.
 4. Their bodies tend to lose too much water to their environment.
-

A

In which situation would passive transport **not** use a transport protein for entry into a cell?

1. water flowing into a hypertonic environment
2. glucose being absorbed from the blood
3. an ion flowing into a nerve cell to create

- an electrical potential
4. oxygen moving into a cell after oxygen deprivation
-

D

Critical Thinking Questions

Discuss why the following affect the rate of diffusion: molecular size, temperature, solution density, and the distance that must be traveled.

Heavy molecules move more slowly than lighter ones. It takes more energy in the medium to move them along. Increasing or decreasing temperature increases or decreases the energy in the medium, affecting molecular movement. The denser a solution is, the harder it is for molecules to move through it, causing diffusion to slow down due to friction. Living cells require a steady supply of nutrients and a steady rate of waste removal. If the distance these substances need to travel is too great, diffusion cannot move nutrients and waste materials efficiently to sustain life.

Why does water move through a membrane?

Water moves through a membrane in osmosis because there is a concentration gradient across the membrane of solute and solvent. The solute cannot effectively move to balance the concentration on both sides of the membrane, so water moves to achieve this balance.

Both of the regular intravenous solutions administered in medicine, normal saline and lactated Ringer's solution, are isotonic. Why is this important?

Injection of isotonic solutions ensures that there will be no perturbation of the osmotic balance, and no water taken from tissues or added to them from the blood.

Describe two ways that decreasing temperature would affect the rate of diffusion of molecules across a cell's plasma membrane.

Decreasing temperature will decrease the kinetic energy in the system. A lower temperature means less energy in the molecules, so they will move at a slower speed.

Lowering temperature also decreases the kinetic energy of the molecules in the plasma membrane, compressing them together. This increases the density of the plasma membrane, which slows diffusion into the cell.

A cell develops a mutation in its potassium channels that prevents the ions from leaving the cell. If the cell's aquaporins are still active, what will happen to the cell? Be sure to describe the tonicity and osmolarity of the cell.

Without functional potassium channels, the potassium ions that are pumped into the cell will accumulate. This increases the osmolarity inside the cell, creating a hypotonic solution. Since the plasma membrane is still selectively permeable to water by the aquaporins, water will flow into the cell. If the potassium concentration is high enough, enough water will eventually flow into the cell to lyse it.

Glossary

aquaporin

channel protein that allows water through the membrane at a very high rate

carrier protein

membrane protein that moves a substance across the plasma membrane by changing its own shape

channel protein

membrane protein that allows a substance to pass through its hollow core across the plasma membrane

concentration gradient

area of high concentration adjacent to an area of low concentration

diffusion

passive transport process of low-molecular weight material according to its concentration gradient

facilitated transport

process by which material moves down a concentration gradient (from high to low concentration) using integral membrane proteins

hypertonic

situation in which extracellular fluid has a higher osmolarity than the fluid inside the cell, resulting in water moving out of the cell

hypotonic

situation in which extracellular fluid has a lower osmolarity than the fluid inside the

cell, resulting in water moving into the cell

isotonic

situation in which the extracellular fluid has the same osmolarity as the fluid inside the cell, resulting in no net water movement into or out of the cell

osmolarity

total amount of substances dissolved in a specific amount of solution

osmosis

transport of water through a semipermeable membrane according to the water's concentration gradient across the membrane that results from the presence of solute that cannot pass through the membrane

passive transport

method of transporting material through a membrane that does not require energy

plasmolysis

detaching the cell membrane from the cell wall and constricting the cell membrane when a plant cell is in a hypertonic solution

selectively permeable

membrane characteristic that allows some substances through

solute

substance dissolved in a liquid to form a solution

tonicity

amount of solute in a solution

transport protein

membrane protein that facilitates a substance's passage across a membrane by binding it

Active Transport

By the end of this section, you will be able to do the following:

- Understand how electrochemical gradients affect ions
- Distinguish between primary active transport and secondary active transport

Active transport mechanisms require the cell's energy, usually in the form of adenosine triphosphate (ATP). If a substance must move into the cell against its concentration gradient—that is, if the substance's concentration inside the cell is greater than its concentration in the extracellular fluid (and vice versa)—the cell must use energy to move the substance. Some active transport mechanisms move small-molecular weight materials, such as ions, through the membrane. Other mechanisms transport much larger molecules.

A uniporter carries one molecule or ion. A symporter carries two different molecules or ions, both in the same direction. An antiporter also carries two different molecules or ions, but in different directions. (credit: modification of work by “Lupask”/Wikimedia Commons)

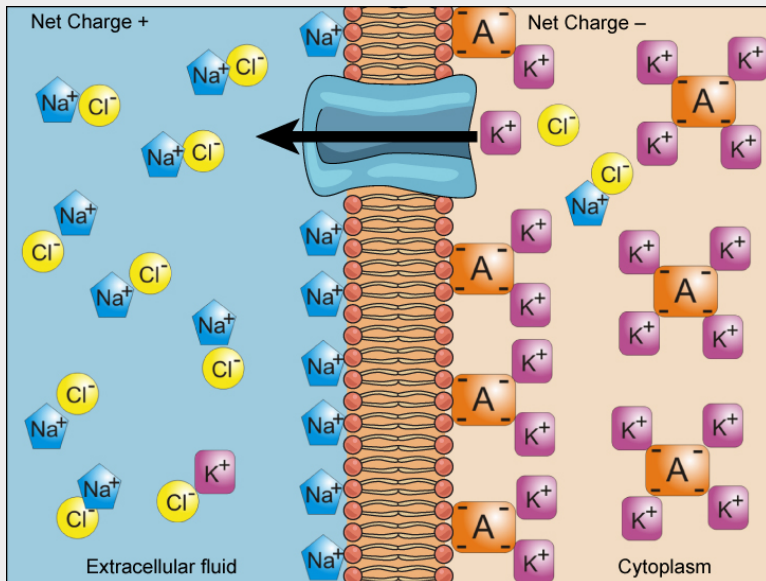
Electrochemical Gradient

We have discussed simple concentration gradients—

a substance's differential concentrations across a space or a membrane—but in living systems, gradients are more complex. Because ions move into and out of cells and because cells contain proteins that do not move across the membrane and are mostly negatively charged, there is also an electrical gradient, a difference of charge, across the plasma membrane. The interior of living cells is electrically negative with respect to the extracellular fluid in which they are bathed, and at the same time, cells have higher concentrations of potassium (K^+) and lower concentrations of sodium (Na^+) than the extracellular fluid. Thus in a living cell, the concentration gradient of Na^+ tends to drive it into the cell, and its electrical gradient (a positive ion) also drives it inward to the negatively charged interior. However, the situation is more complex for other elements such as potassium. The electrical gradient of K^+ , a positive ion, also drives it into the cell, but the concentration gradient of K^+ drives K^+ *out* of the cell ([\[link\]](#)). We call the combined concentration gradient and electrical charge that affects an ion its **electrochemical gradient**.

Visual Connection

Electrochemical gradients arise from the combined effects of concentration gradients and electrical gradients. Structures labeled A represent proteins. (credit: “Synaptitude”/Wikimedia Commons)



Injecting a potassium solution into a person's blood is lethal. This is how capital punishment and euthanasia subjects die. Why do you think a potassium solution injection is lethal?

Moving Against a Gradient

To move substances against a concentration or electrochemical gradient, the cell must use energy. This energy comes from ATP generated through the cell's metabolism. Active transport mechanisms, or **pumps**, work against electrochemical gradients. Small substances constantly pass through plasma membranes. Active transport maintains concentrations of ions and other substances that living cells require in the face of these passive

movements. A cell may spend much of its metabolic energy supply maintaining these processes. (A red blood cell uses most of its metabolic energy to maintain the imbalance between exterior and interior sodium and potassium levels that the cell requires.) Because active transport mechanisms depend on a cell's metabolism for energy, they are sensitive to many metabolic poisons that interfere with the ATP supply.

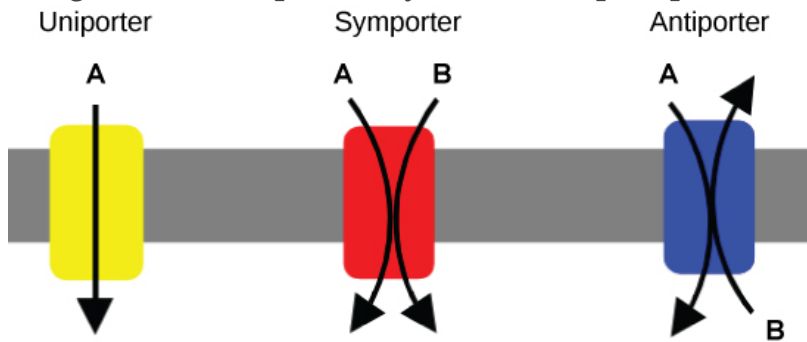
Two mechanisms exist for transporting small-molecular weight material and small molecules.

Primary active transport moves ions across a membrane and creates a difference in charge across that membrane, which is directly dependent on ATP. **Secondary active transport** does not directly require ATP: instead, it is the movement of material due to the electrochemical gradient established by primary active transport.

Carrier Proteins for Active Transport

An important membrane adaption for active transport is the presence of specific carrier proteins or pumps to facilitate movement: there are three protein types or **transporters** ([\[link\]](#)). A **uniporter** carries one specific ion or molecule. A **symporter** carries two different ions or molecules, both in the same direction. An **antiporter** also carries two different ions or molecules, but in different directions. All of these transporters can also

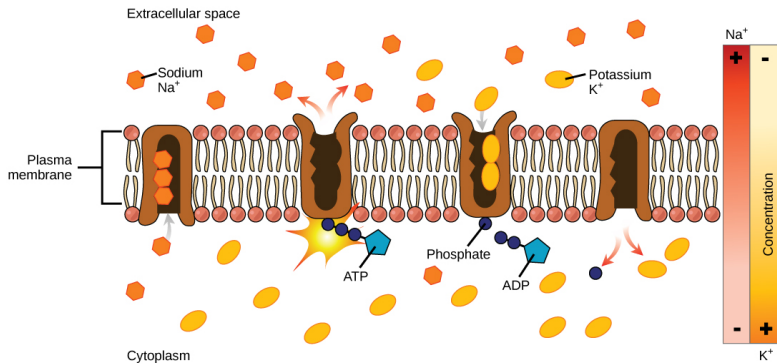
transport small, uncharged organic molecules like glucose. These three types of carrier proteins are also in facilitated diffusion, but they do not require ATP to work in that process. Some examples of pumps for active transport are $\text{Na}^+ - \text{K}^+$ ATPase, which carries sodium and potassium ions, and $\text{H}^+ - \text{K}^+$ ATPase, which carries hydrogen and potassium ions. Both of these are antiporter carrier proteins. Two other carrier proteins are Ca^{2+} ATPase and H^+ ATPase, which carry only calcium and only hydrogen ions, respectively. Both are pumps.



Primary active transport moves ions across a membrane, creating an electrochemical gradient (electrogenic transport). (credit: modification of work by Mariana Ruiz Villareal)

Primary Active Transport

The primary active transport that functions with the active transport of sodium and potassium allows secondary active transport to occur. The second transport method is still active because it depends on using energy as does primary transport ([\[link\]](#)).



One of the most important pumps in animal cells is the sodium-potassium pump ($\text{Na}^+ - \text{K}^+ \text{ ATPase}$), which maintains the electrochemical gradient (and the correct concentrations of Na^+ and K^+) in living cells. The sodium-potassium pump moves K^+ into the cell while moving Na^+ out at the same time, at a ratio of three Na^+ for every two K^+ ions moved in. The $\text{Na}^+ - \text{K}^+ \text{ ATPase}$ exists in two forms, depending on its orientation to the cell's interior or exterior and its affinity for either sodium or potassium ions. The process consists of the following six steps.

1. With the enzyme oriented towards the cell's interior, the carrier has a high affinity for sodium ions. Three ions bind to the protein.
2. The protein carrier hydrolyzes ATP and a low-energy phosphate group attaches to it.
3. As a result, the carrier changes shape and reorients itself towards the membrane's exterior. The protein's affinity for sodium decreases and the three sodium ions leave the

carrier.

4. The shape change increases the carrier's affinity for potassium ions, and two such ions attach to the protein. Subsequently, the low-energy phosphate group detaches from the carrier.
5. With the phosphate group removed and potassium ions attached, the carrier protein repositions itself towards the cell's interior.
6. The carrier protein, in its new configuration, has a decreased affinity for potassium, and the two ions move into the cytoplasm. The protein now has a higher affinity for sodium ions, and the process starts again.

Several things have happened as a result of this process. At this point, there are more sodium ions outside the cell than inside and more potassium ions inside than out. For every three sodium ions that move out, two potassium ions move in. This results in the interior being slightly more negative relative to the exterior. This difference in charge is important in creating the conditions necessary for the secondary process. The sodium-potassium pump is, therefore, an **electrogenic pump** (a pump that creates a charge imbalance), creating an electrical imbalance across the membrane and contributing to the membrane potential.

Link to Learning

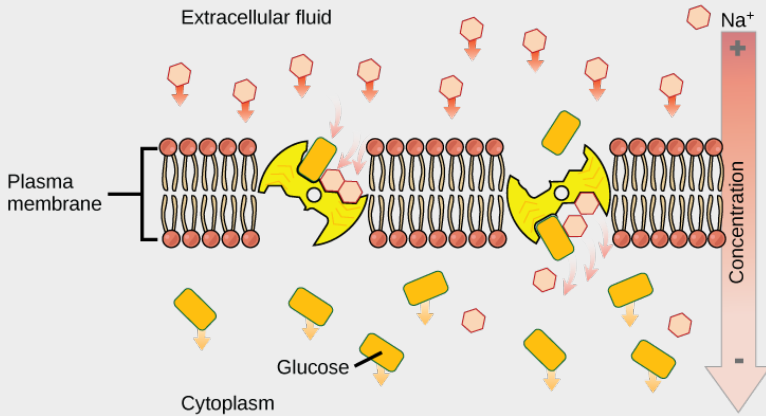
Watch this [video](#) to see an active transport simulation in a sodium-potassium ATPase.

Secondary Active Transport (Co-transport)

Secondary active transport brings sodium ions, and possibly other compounds, into the cell. As sodium ion concentrations build outside of the plasma membrane because of the primary active transport process, this creates an electrochemical gradient. If a channel protein exists and is open, the sodium ions will pull through the membrane. This movement transports other substances that can attach themselves to the transport protein through the membrane ([\[link\]](#)). Many amino acids, as well as glucose, enter a cell this way. This secondary process also stores high-energy hydrogen ions in the mitochondria of plant and animal cells in order to produce ATP. The potential energy that accumulates in the stored hydrogen ions translates into kinetic energy as the ions surge through the channel protein ATP synthase, and that energy then converts ADP into ATP.

Visual Connection

An electrochemical gradient, which primary active transport creates, can move other substances against their concentration gradients, a process scientists call co-transport or secondary active transport. (credit: modification of work by Mariana Ruiz Villareal)



If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

Section Summary

The combined gradient that affects an ion includes its concentration gradient and its electrical gradient. A positive ion, for example, might diffuse into a new area, down its concentration gradient, but if it is

diffusing into an area of net positive charge, its electrical gradient hampers its diffusion. When dealing with ions in aqueous solutions, one must consider electrochemical and concentration gradient combinations, rather than just the concentration gradient alone. Living cells need certain substances that exist inside the cell in concentrations greater than they exist in the extracellular space. Moving substances up their electrochemical gradients requires energy from the cell. Active transport uses energy stored in ATP to fuel this transport. Active transport of small molecular-sized materials uses integral proteins in the cell membrane to move the materials. These proteins are analogous to pumps. Some pumps, which carry out primary active transport, couple directly with ATP to drive their action. In co-transport (or secondary active transport), energy from primary transport can move another substance into the cell and up its concentration gradient.

Visual Connection Questions

[\[link\]](#) Injecting a potassium solution into a person's blood is lethal. Capital punishment and euthanasia utilize this method in their subjects. Why do you think a potassium solution injection is lethal?

[\[link\]](#) Cells typically have a high concentration of potassium in the cytoplasm and are bathed in a high concentration of sodium. Injection of potassium dissipates this electrochemical gradient. In heart muscle, the sodium/potassium potential is responsible for transmitting the signal that causes the muscle to contract. When this potential is dissipated, the signal can't be transmitted, and the heart stops beating. Potassium injections are also used to stop the heart from beating during surgery.

[\[link\]](#) If the pH outside the cell decreases, would you expect the amount of amino acids transported into the cell to increase or decrease?

[\[link\]](#) A decrease in pH means an increase in positively charged H^+ ions, and an increase in the electrical gradient across the membrane. The transport of amino acids into the cell will increase.

Review Questions

Active transport must function continuously because _____.

1. plasma membranes wear out
 2. not all membranes are amphiphilic
 3. facilitated transport opposes active transport
 4. diffusion is constantly moving solutes in opposite directions
-

D

How does the sodium-potassium pump make the interior of the cell negatively charged?

1. by expelling anions
 2. by pulling in anions
 3. by expelling more cations than are taken in
 4. by taking in and expelling an equal number of cations
-

C

What is the combination of an electrical gradient and a concentration gradient called?

1. potential gradient
2. electrical potential

3. concentration potential
 4. electrochemical gradient
-

D

Critical Thinking Questions

Where does the cell get energy for active transport processes?

The cell harvests energy from ATP produced by its own metabolism to power active transport processes, such as the activity of pumps.

How does the sodium-potassium pump contribute to the net negative charge of the interior of the cell?

The sodium-potassium pump forces out three (positive) Na^+ ions for every two (positive) K^+ ions it pumps in, thus the cell loses a positive charge at every cycle of the pump.

Glucose from digested food enters intestinal

epithelial cells by active transport. Why would intestinal cells use active transport when most body cells use facilitated diffusion?

Intestinal epithelial cells use active transport to fulfill their specific role as the cells that transfer glucose from the digested food to the bloodstream. Intestinal cells are exposed to an environment with fluctuating glucose levels. Immediately after eating, glucose in the gut lumen will be high, and could accumulate in intestinal cells by diffusion. However, when the gut lumen is empty, glucose levels are higher in the intestinal cells. If glucose moved by facilitated diffusion, this would cause glucose to flow back out of the intestinal cells and into the gut. Active transport proteins ensure that glucose moves into the intestinal cells, and cannot move back into the gut. It also ensures that glucose transport continues to occur even if high levels of glucose are already present in the intestinal cells. This maximizes the amount of energy the body can harvest from food.

The sodium/calcium exchanger (NCX) transports sodium into and calcium out of cardiac muscle cells. Describe why this transporter is classified as secondary active transport.

The NCX moves sodium down its electrochemical gradient into the cell. Since sodium's electrochemical gradient is created by the Na^+/K^+ pump, a transport pump that requires ATP hydrolysis to establish the gradient, the NCX is a secondary active transport process.

Glossary

active transport

method of transporting material that requires energy

antiporter

transporter that carries two ions or small molecules in different directions

electrochemical gradient

a combined electrical and chemical force that produces a gradient

electrogenic pump

pump that creates a charge imbalance

primary active transport

active transport that moves ions or small molecules across a membrane and may create a difference in charge across that membrane

pump

active transport mechanism that works against electrochemical gradients

secondary active transport

movement of material that results from primary active transport to the electrochemical gradient

symporter

transporter that carries two different ions or small molecules, both in the same direction

transporter

specific carrier proteins or pumps that facilitate movement

uniporter

transporter that carries one specific ion or molecule

Bulk Transport

By the end of this section, you will be able to do the following:

- Describe endocytosis, including phagocytosis, pinocytosis, and receptor-mediated endocytosis
- Understand the process of exocytosis

In addition to moving small ions and molecules through the membrane, cells also need to remove and take in larger molecules and particles (see [\[link\]](#) for examples). Some cells are even capable of engulfing entire unicellular microorganisms. You might have correctly hypothesized that when a cell uptakes and releases large particles, it requires energy. A large particle, however, cannot pass through the membrane, even with energy that the cell supplies.

In phagocytosis, the cell membrane surrounds the particle and engulfs it. (credit: modification of work by Mariana Ruiz Villareal) In pinocytosis, the cell membrane invaginates, surrounds a small volume of fluid, and pinches off. (credit: modification of work by Mariana Ruiz Villareal) In receptor-mediated endocytosis, the cell's uptake of substances targets a single type of substance that binds to the receptor on the cell membrane's external surface. (credit: modification of work by Mariana Ruiz Villareal)

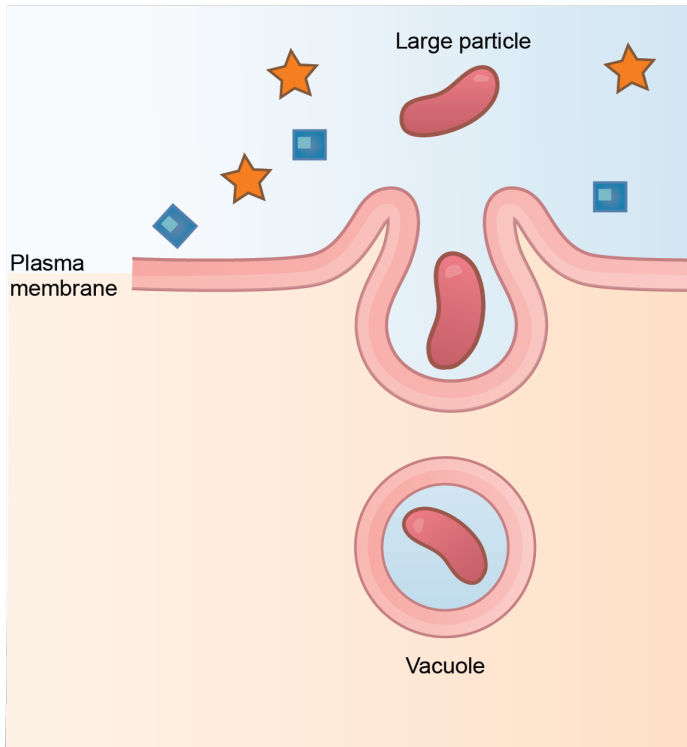
Endocytosis

Endocytosis is a type of active transport that moves particles, such as large molecules, parts of cells, and even whole cells, into a cell. There are different endocytosis variations, but all share a common characteristic: the cell's plasma membrane invaginates, forming a pocket around the target particle. The pocket pinches off, resulting in the particle containing itself in a newly created intracellular vesicle formed from the plasma membrane.

Phagocytosis

Phagocytosis (the condition of “cell eating”) is the process by which a cell takes in large particles, such as other cells or relatively large particles. For example, when microorganisms invade the human body, a type of white blood cell, a neutrophil, will remove the invaders through this process, surrounding and engulfing the microorganism, which the neutrophil then destroys ([\[link\]](#)).

Phagocytosis



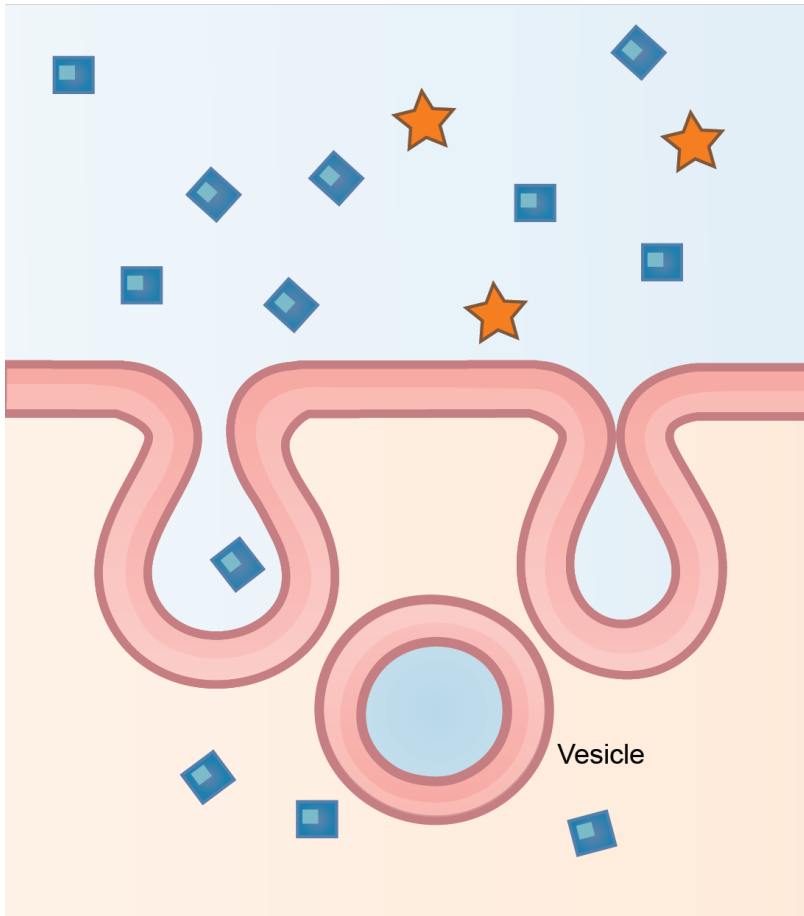
In preparation for phagocytosis, a portion of the plasma membrane's inward-facing surface becomes coated with the protein **clathrin**, which stabilizes this membrane's section. The membrane's coated portion then extends from the cell's body and surrounds the particle, eventually enclosing it. Once the vesicle containing the particle is enclosed within the cell, the clathrin disengages from the membrane and the vesicle merges with a lysosome for breaking down the material in the newly formed compartment (endosome). When accessible nutrients from the vesicular contents' degradation have been extracted, the newly formed endosome

merges with the plasma membrane and releases its contents into the extracellular fluid. The endosomal membrane again becomes part of the plasma membrane.

Pinocytosis

A variation of endocytosis is **pinocytosis**. This literally means “cell drinking”. Discovered by Warren Lewis in 1929, this American embryologist and cell biologist described a process whereby he assumed that the cell was purposefully taking in extracellular fluid. In reality, this is a process that takes in molecules, including water, which the cell needs from the extracellular fluid. Pinocytosis results in a much smaller vesicle than does phagocytosis, and the vesicle does not need to merge with a lysosome ([\[link\]](#)).

Pinocytosis



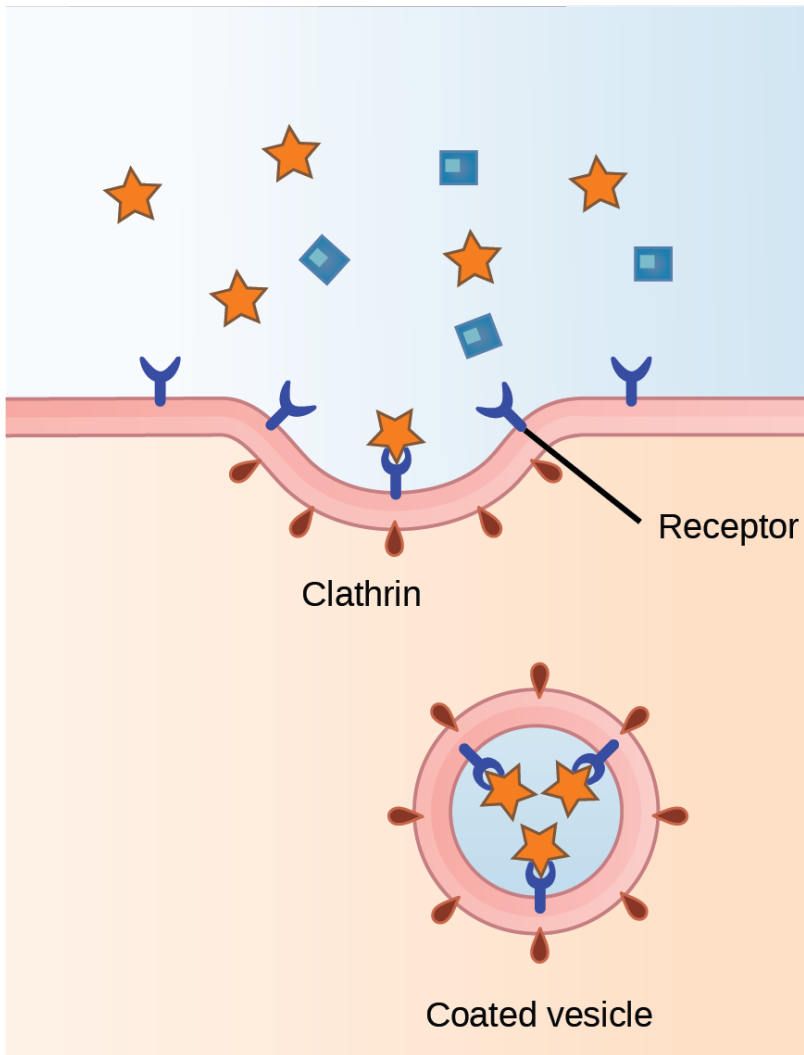
A variation of pinocytosis is **potocytosis**. This process uses a coating protein, **caveolin**, on the plasma membrane's cytoplasmic side, which performs a similar function to clathrin. The cavities in the plasma membrane that form the vacuoles have membrane receptors and lipid rafts in addition to caveolin. The vacuoles or vesicles formed in caveolae (singular caveola) are smaller than those in pinocytosis. Potocytosis brings small molecules into

the cell and transports them through the cell for their release on the other side, a process we call transcytosis.

Receptor-mediated Endocytosis

A targeted variation of endocytosis employs receptor proteins in the plasma membrane that have a specific binding affinity for certain substances ([link](#)).

Receptor-mediated endocytosis



In **receptor-mediated endocytosis**, as in phagocytosis, clathrin attaches to the plasma membrane's cytoplasmic side. If a compound's uptake is dependent on receptor-mediated endocytosis and the process is ineffective, the

material will not be removed from the tissue fluids or blood. Instead, it will stay in those fluids and increase in concentration. The failure of receptor-mediated endocytosis causes some human diseases. For example, receptor mediated endocytosis removes low density lipoprotein or LDL (or "bad" cholesterol) from the blood. In the human genetic disease familial hypercholesterolemia, the LDL receptors are defective or missing entirely. People with this condition have life-threatening levels of cholesterol in their blood, because their cells cannot clear LDL particles.

Although receptor-mediated endocytosis is designed to bring specific substances that are normally in the extracellular fluid into the cell, other substances may gain entry into the cell at the same site. Flu viruses, diphtheria, and cholera toxin all have sites that cross-react with normal receptor-binding sites and gain entry into cells.

Link to Learning

See receptor-mediated endocytosis in action, and click on different [parts](#) for a focused animation.

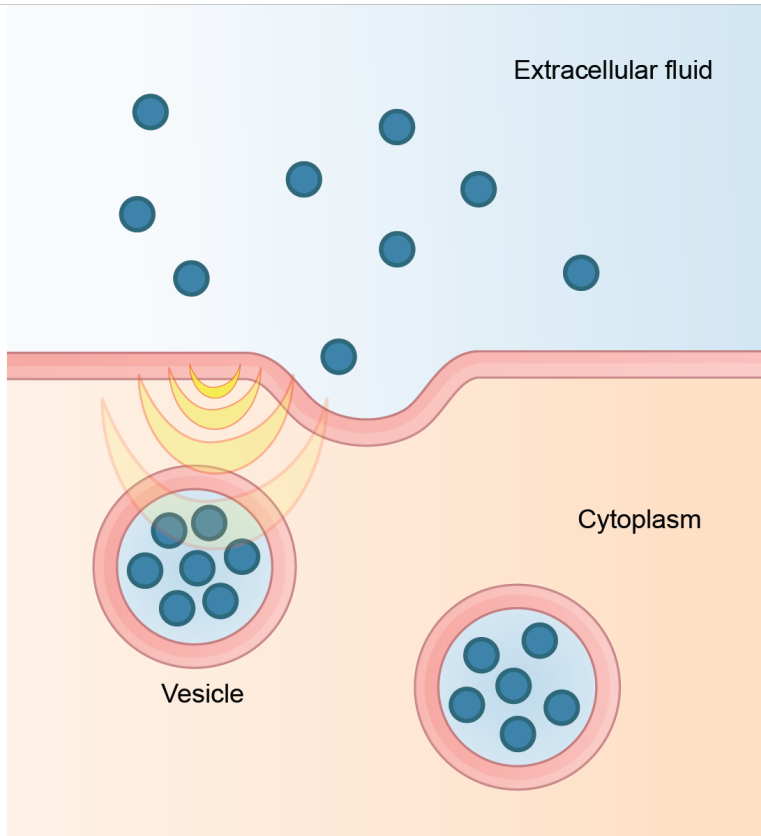
In exocytosis, vesicles containing substances fuse with the plasma membrane. The contents then

release to the cell's exterior. (credit: modification of work by Mariana Ruiz Villareal)

Exocytosis

The reverse process of moving material into a cell is the process of exocytosis. **Exocytosis** is the opposite of the processes we discussed above in that its purpose is to expel material from the cell into the extracellular fluid. Waste material is enveloped in a membrane and fuses with the plasma membrane's interior. This fusion opens the membranous envelope on the cell's exterior, and the waste material expels into the extracellular space ([\[link\]](#)). Other examples of cells releasing molecules via exocytosis include extracellular matrix protein secretion and neurotransmitter secretion into the synaptic cleft by synaptic vesicles.

Exocytosis



**Methods of
Transport,
Energy
Requirements,
and Types of
Transported**

Material Transport Method	Active/Passive	Material Transported
Diffusion	Passive	Small-molecular weight material
Osmosis	Passive	Water
Facilitated transport/diffusion	Passive	Sodium, potassium, calcium, glucose
Primary active transport	Active	Sodium, potassium, calcium
Secondary active transport	Active	Amino acids, lactose
Phagocytosis	Active	Large macromolecules, whole cells, or cellular structures
Pinocytosis and potocytosis	Active	Small molecules (liquids/water)
Receptor-mediated endocytosis	Active	Large quantities of macromolecules

Section Summary

Active transport methods require directly using ATP to fuel the transport. In a process scientists call

phagocytosis, other cells can engulf large particles, such as macromolecules, cell parts, or whole cells. In phagocytosis, a portion of the membrane invaginates and flows around the particle, eventually pinching off and leaving the particle entirely enclosed by a plasma membrane's envelope. The cell breaks down vesicle contents, with the particles either used as food or dispatched. Pinocytosis is a similar process on a smaller scale. The plasma membrane invaginates and pinches off, producing a small envelope of fluid from outside the cell. Pinocytosis imports substances that the cell needs from the extracellular fluid. The cell expels waste in a similar but reverse manner. It pushes a membranous vacuole to the plasma membrane, allowing the vacuole to fuse with the membrane and incorporate itself into the membrane structure, releasing its contents to the exterior.

Review Questions

What happens to the membrane of a vesicle after exocytosis?

1. It leaves the cell.
2. It is disassembled by the cell.
3. It fuses with and becomes part of the plasma membrane.

4. It is used again in another exocytosis event.
-

C

Which transport mechanism can bring whole cells into a cell?

1. pinocytosis
 2. phagocytosis
 3. facilitated transport
 4. primary active transport
-

B

In what important way does receptor-mediated endocytosis differ from phagocytosis?

1. It transports only small amounts of fluid.
 2. It does not involve the pinching off of membrane.
 3. It brings in only a specifically targeted substance.
 4. It brings substances into the cell, while phagocytosis removes substances.
-

C

Many viruses enter host cells through receptor-mediated endocytosis. What is an advantage of this entry strategy?

1. The virus directly enters the cytoplasm of the cell.
 2. The virus is protected from recognition by white blood cells.
 3. The virus only enters its target host cell type.
 4. The virus can directly inject its genome into the cell's nucleus.
-

C

Which of the following organelles relies on exocytosis to complete its function?

1. Golgi apparatus
 2. vacuole
 3. mitochondria
 4. endoplasmic reticulum
-

A

Imagine a cell can perform exocytosis, but only minimal endocytosis. What would happen to the cell?

1. The cell would secrete all its intracellular proteins.
2. The plasma membrane would increase in size over time.
3. The cell would stop expressing integral receptor proteins in its plasma membrane.
4. The cell would lyse.

B

Critical Thinking Questions

Why is it important that there are different types of proteins in plasma membranes for the transport of materials into and out of a cell?

The proteins allow a cell to select what compound will be transported, meeting the needs of the cell and not bringing in anything else.

Why do ions have a difficult time getting through plasma membranes despite their small size?

Ions are charged, and consequently, they are hydrophilic and cannot associate with the lipid portion of the membrane. Ions must be transported by carrier proteins or ion channels.

Glossary

caveolin

protein that coats the plasma membrane's cytoplasmic side and participates in the liquid uptake process by potocytosis

clathrin

protein that coats the plasma membrane's inward-facing surface and assists in forming specialized structures, like coated pits, for phagocytosis

endocytosis

type of active transport that moves substances, including fluids and particles, into a cell

exocytosis

process of passing bulk material out of a cell

pinocytosis

a variation of endocytosis that imports macromolecules that the cell needs from the extracellular fluid

potocytosis

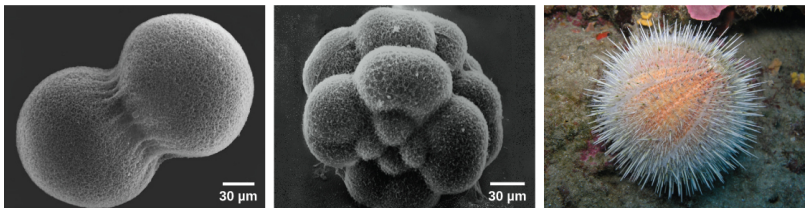
variation of pinocytosis that uses a different coating protein (caveolin) on the plasma membrane's cytoplasmic side

receptor-mediated endocytosis

variation of endocytosis that involves using specific binding proteins in the plasma membrane for specific molecules or particles, and clathrin-coated pits that become clathrin-coated vesicles

Introduction

class = "introduction" A sea urchin begins life as a single diploid cell (zygote) that (a) divides through cell division to form two genetically identical daughter cells, visible here through scanning electron microscopy (SEM). After four rounds of cell division, (b) there are 16 cells, as seen in this SEM image. After many rounds of cell division, the individual develops into a complex, multicellular organism, as seen in this (c) mature sea urchin. (credit a: modification of work by Evelyn Spiegel, Louisa Howard; credit b: modification of work by Evelyn Spiegel, Louisa Howard; credit c: modification of work by Marco Busdraghi; scale-bar data from Matt Russell)



A human, like every sexually reproducing organism, begins life as a fertilized egg (embryo) or **zygote**. In our species, billions of cell divisions subsequently must occur in a controlled manner in order to produce a complex, multicellular human comprising trillions of cells. Thus, the original single-celled zygote is literally the ancestor of all cells in the body. However, once a human is fully grown, cell reproduction is still necessary to repair and regenerate tissues, and sometimes to increase our

size! In fact, all multicellular organisms use cell division for growth and the maintenance and repair of cells and tissues. Cell division is closely regulated, and the occasional failure of this regulation can have life-threatening consequences. Single-celled organisms may also use cell division as their method of reproduction.

Cell Division

By the end of this section, you will be able to do the following:

- Describe the structure of prokaryotic and eukaryotic genomes
- Distinguish between chromosomes, genes, and traits
- Describe the mechanisms of chromosome compaction

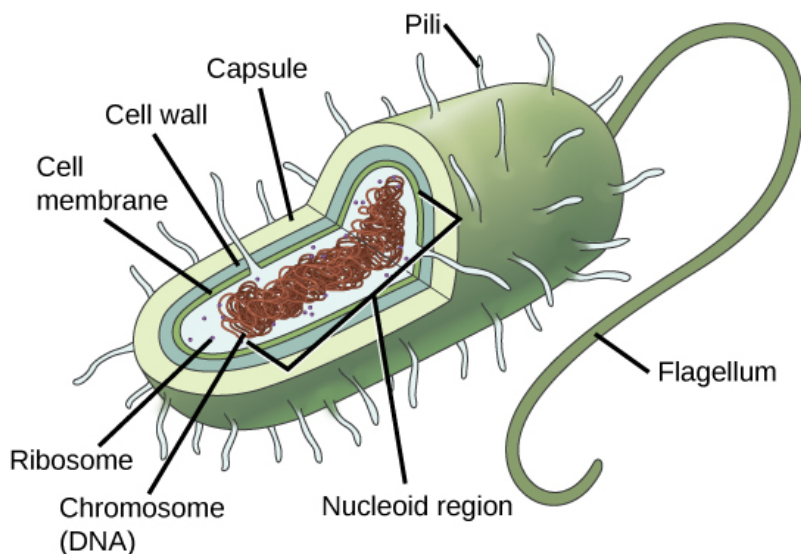
The continuity of life from one cell to another has its foundation in the reproduction of cells by way of the cell cycle. The cell cycle is an orderly sequence of events that describes the stages of a cell's life from the division of a single parent cell to the production of two new genetically identical daughter cells.

Prokaryotes, including both Bacteria and Archaea, have a single, circular chromosome located in a central region called the nucleoid. There are 23 pairs of homologous chromosomes in a female human somatic cell. The condensed chromosomes are viewed within the nucleus (top), removed from a cell during mitosis (also called karyokinesis or nuclear division) and spread out on a slide (right), and artificially arranged according to length (left); an arrangement like this is called a karyotype. In this image, the chromosomes were exposed to fluorescent stains for differentiation of the different chromosomes. A method of staining called

“*chromosome painting*” employs fluorescent dyes that highlight chromosomes in different colors. (credit: National Human Genome Project/NIH)

Genomic DNA

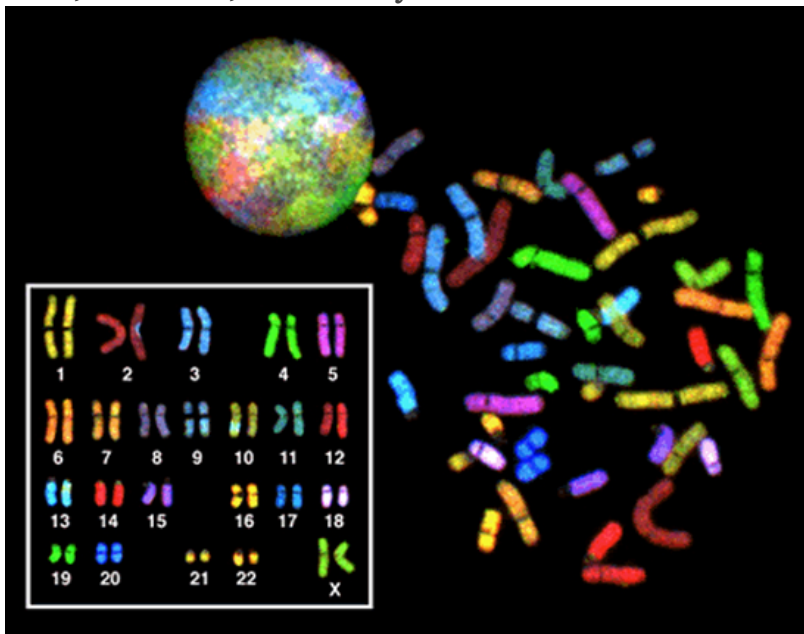
Before discussing the steps a cell must undertake to replicate and divide its DNA, a deeper understanding of the structure and function of a cell’s genetic information is necessary. A cell’s DNA, packaged as a double-stranded DNA molecule, is called its **genome**. In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle ([\[link\]](#)). The region in the cell containing this genetic material is called a nucleoid. Some prokaryotes also have smaller loops of DNA called plasmids that are not essential for normal growth. Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. *Antibiotic resistance* is one trait that often spreads through a bacterial colony through plasmid exchange from resistant donors to recipient cells.



In eukaryotes, the genome consists of several double-stranded linear DNA molecules ([\[link\]](#)). Each species of eukaryotes has a characteristic number of chromosomes in the nuclei of its cells. Human body (somatic) cells have 46 chromosomes, while human **gametes** (sperm or eggs) have 23 chromosomes each. A typical body cell contains two matched or homologous sets of chromosomes (one set from each biological parent)—a configuration known as **diploid**. (Note: The letter n is used to represent a single set of chromosomes; therefore, a diploid organism is designated $2n$.) Human cells that contain one set of chromosomes are called gametes, or sex cells; these are eggs and sperm, and are designated $1n$, or **haploid**.

Upon fertilization, each gamete contributes one set of chromosomes, creating a diploid cell containing

matched pairs of chromosomes called **homologous** (“same knowledge”) **chromosomes**. Homologous chromosomes are the same length and have specific nucleotide segments called **genes** in exactly the same location, or **locus**. Genes, the functional units of chromosomes, determine specific characteristics by coding for specific proteins. Traits are the variations of those characteristics. For example, hair color is a characteristic with traits that are blonde, brown, or black, and many colors in between.



Each copy of a homologous pair of chromosomes originates from a different parent; therefore, the different genes (alleles) themselves are not identical, although they code for the same traits such as “hair color.” The variation of individuals within a species is due to the specific combination of the genes

inherited from both parents. Even a slightly altered sequence of nucleotides within a gene can result in an alternative trait. For example, there are three possible gene sequences on the human chromosome that code for blood type: sequence A, sequence B, and sequence O. Because all diploid human cells have two copies of the chromosome that determines blood type, the blood type (the trait) is determined by the two alleles of the marker gene that are inherited. It is possible to have two copies of the same gene sequence on both homologous chromosomes, with one on each (for example, AA, BB, or OO), or two different sequences, such as AB, AO, or BO.

Apparently minor variations of traits, such as blood type, eye color, and handedness, contribute to the natural variation found within a species, but even though they seem minor, these traits may be connected with the expression of other traits as of yet unknown. However, if the entire DNA sequence from any pair of human homologous chromosomes is compared, the difference is much less than one percent. The sex chromosomes, X and Y, are the single exception to the rule of homologous chromosome uniformity: Other than a small amount of homology that is necessary to accurately produce gametes, the genes found on the X and Y chromosomes are different.

Double-stranded DNA wraps around histone proteins to form nucleosomes that create the

appearance of “beads on a string.” The nucleosomes are coiled into a 30-nm chromatin fiber. When a cell undergoes mitosis, the chromosomes condense even further.

Eukaryotic Chromosomal Structure and Compaction

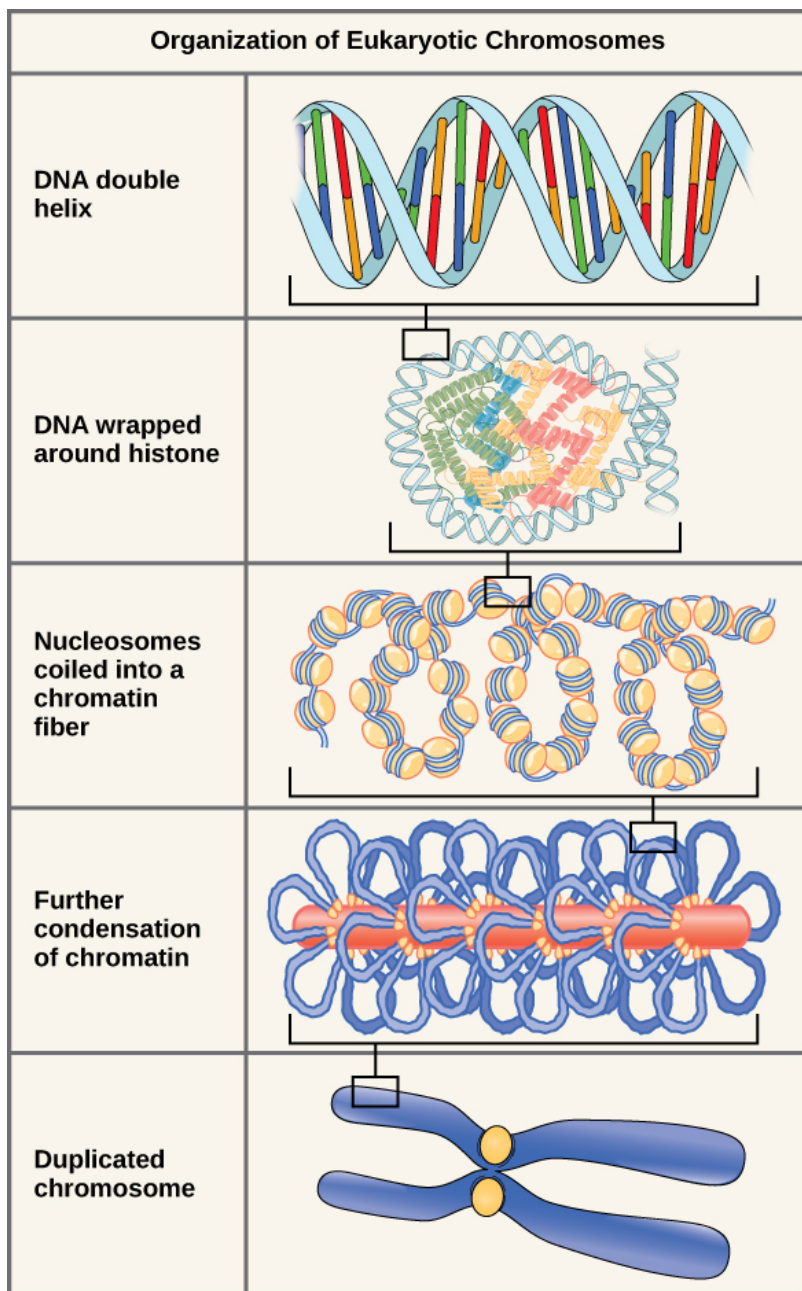
If the DNA from all 46 chromosomes in a human cell nucleus were laid out end-to-end, it would measure approximately two meters; however, its diameter would be only 2 nm! Considering that the size of a typical human cell is about 10 μm (100,000 cells lined up to equal one meter), DNA must be *tightly packaged* to fit in the cell’s nucleus. At the same time, it must also be readily accessible for the genes to be expressed. For this reason, the long strands of DNA are condensed into compact chromosomes during certain stages of the cell cycle. There are a number of ways that chromosomes are compacted.

In the first level of compaction, short stretches of the DNA double helix wrap around a core of eight **histone proteins** at regular intervals along the entire length of the chromosome ([\[link\]](#)). The DNA-histone complex is called chromatin. The beadlike, histone DNA complex is called a **nucleosome**, and DNA connecting the nucleosomes is called linker DNA. A DNA molecule in this form is about seven times shorter than the double helix without the

histones, and the beads are about 10 nm in diameter, in contrast with the 2-nm diameter of a DNA double helix.

The second level of compaction occurs as the nucleosomes and the linker DNA between them coil into a 30-nm chromatin fiber. This coiling further *condenses* the chromosome so that it is now about 50 times shorter than the extended form.

In the third level of compaction, a variety of *fibrous proteins* is used to “pack the chromatin.” These fibrous proteins also ensure that each chromosome in a non-dividing cell occupies a particular area of the nucleus that does not overlap with that of any other chromosome (see the top image in [\[link\]](#)).



DNA replicates in the S phase of interphase, which technically is not a part of mitosis, but must always

precede it. After replication, the chromosomes are composed of two linked sister **chromatids**. When fully compact, the pairs of identically packed chromosomes are bound to each other by cohesin proteins. The connection between the sister chromatids is closest in a region called the **centromere**. The conjoined sister chromatids, with a diameter of about 1 μm , are visible under a light microscope. The centromeric region is highly condensed and thus will appear as a constricted area.

Link to Learning

This animation illustrates the different levels of chromosome packing.

https://www.openstax.org/l/Packaged_DNA

Section Summary

Prokaryotes have a single circular chromosome composed of double-stranded DNA, whereas eukaryotes have multiple, linear chromosomes composed of chromatin wrapped around histones, all of which are surrounded by a nuclear membrane. The 46 chromosomes of human somatic cells are

composed of 22 pairs of autosomes (matched pairs) and a pair of sex chromosomes, which may or may not be matched. This is the $2n$ or diploid state. Human gametes have 23 chromosomes, or one complete set of chromosomes; a set of chromosomes is complete with either one of the sex chromosomes, X or Y. This is the n or haploid state. Genes are segments of DNA that code for a specific functional molecule (a protein or RNA). An organism's traits are determined by the genes inherited from each parent. Duplicated chromosomes are composed of two sister chromatids. Chromosomes are compacted using a variety of mechanisms during certain stages of the cell cycle. Several classes of protein are involved in the organization and packing of the chromosomal DNA into a highly condensed structure. The condensing complex compacts chromosomes, and the resulting condensed structure is necessary for chromosomal segregation during mitosis.

Review Questions

A diploid cell has_____ the number of chromosomes as a haploid cell.

1. one-fourth
2. half

3. twice
4. four times

C

An organism's traits are determined by the specific combination of inherited ____.

1. cells.
2. genes.
3. proteins.
4. chromatids.

B

The first level of DNA organization in a eukaryotic cell is maintained by which molecule?

1. cohesin
2. condensin
3. chromatin
4. histone

D

Identical copies of chromatin held together by cohesin at the centromere are called ____.

1. histones.
2. nucleosomes.
3. chromatin.
4. sister chromatids.

D

Critical Thinking Questions

Compare and contrast a human somatic cell to a human gamete.

Human somatic cells have 46 chromosomes: 22 pairs and 2 sex chromosomes that may or may not form a pair. This is the $2n$ or diploid condition. Human gametes have 23 chromosomes, one each of 23 unique chromosomes, one of which is a sex chromosome. This is the n or haploid condition.

What is the relationship between a genome, chromosomes, and genes?

The genome consists of the sum total of an organism's chromosomes. Each chromosome contains hundreds and sometimes thousands of genes, segments of DNA that code for a polypeptide or RNA, and a large amount of DNA with no known function.

Eukaryotic chromosomes are thousands of times longer than a typical cell. Explain how chromosomes can fit inside a eukaryotic nucleus.

The DNA double helix is wrapped around histone proteins to form structures called nucleosomes. Nucleosomes and the linker DNA in between them are coiled into a 30-nm fiber. During cell division, chromatin is further condensed by packing proteins.

Glossary

cell cycle

ordered sequence of events through which a cell passes between one cell division and the next

centromere

region at which sister chromatids are bound

together; a constricted area in condensed chromosomes

chromatid

single DNA molecule of two strands of duplicated DNA and associated proteins held together at the centromere

diploid

cell, nucleus, or organism containing two sets of chromosomes ($2n$)

gamete

haploid reproductive cell or sex cell (sperm, pollen grain, or egg)

gene

physical and functional unit of heredity, a sequence of DNA that codes for a protein.

genome

total genetic information of a cell or organism

haploid

cell, nucleus, or organism containing one set of chromosomes (n)

histone

one of several similar, highly conserved, low molecular weight, basic proteins found in the chromatin of all eukaryotic cells; associates with DNA to form nucleosomes

homologous chromosomes

chromosomes of the same morphology with genes in the same location; diploid organisms have pairs of homologous chromosomes (homologs), with each homolog derived from a different parent

locus

position of a gene on a chromosome

nucleosome

subunit of chromatin composed of a short length of DNA wrapped around a core of histone proteins

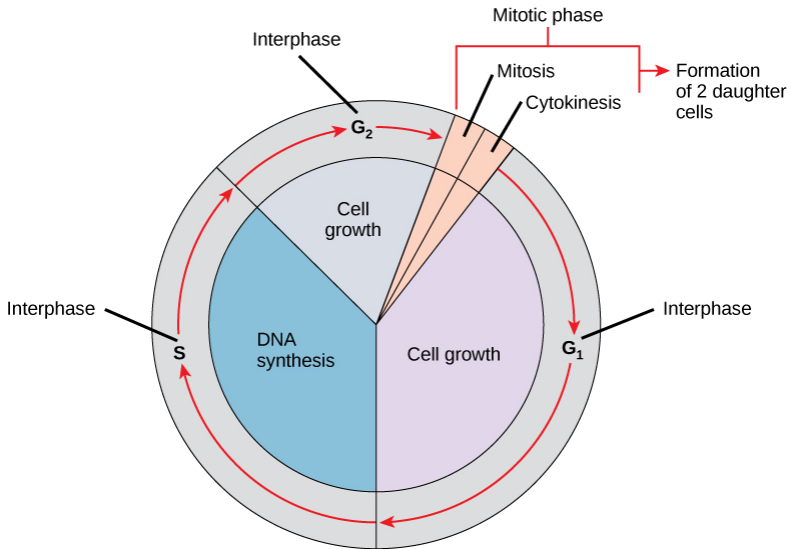
The Cell Cycle

By the end of this section, you will be able to do the following:

- Describe the three stages of interphase
- Discuss the behavior of chromosomes during karyokinesis/mitosis
- Explain how the cytoplasmic content is divided during cytokinesis
- Define the quiescent G₀ phase

The **cell cycle** is an ordered series of events involving cell growth and cell division that produces two new daughter cells. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages of growth, DNA replication, and nuclear and cytoplasmic division that ultimately produces two identical (clone) cells. The cell cycle has two major phases: interphase and the mitotic phase ([\[link\]](#)). During **interphase**, the cell grows and DNA is replicated. During the **mitotic phase**, the replicated DNA and cytoplasmic contents are separated, and the cell cytoplasm is typically partitioned by a third process of the cell cycle called **cytokinesis**. We should note, however, that interphase and mitosis (karyokinesis) may take place without cytokinesis, in which case cells with multiple nuclei (multinucleate cells) are produced. The cell cycle in multicellular organisms consists of interphase and the mitotic phase. During interphase, the cell grows and the nuclear DNA is duplicated.

Interphase is followed by the mitotic phase. During the mitotic phase, the duplicated chromosomes are segregated and distributed into daughter nuclei. Following mitosis, the cytoplasm is usually divided as well by cytokinesis, resulting in two genetically identical daughter cells.



Interphase

During interphase, the cell undergoes normal growth processes while also preparing for cell division. In order for a cell to move from interphase into the mitotic phase, many internal and external conditions must be met. The three stages of interphase are called *G₁*, *S*, and *G₂*.

G₁ Phase (First Gap)

The first stage of interphase is called the **G₁ phase** (first gap) because, from a microscopic point of view, little change is visible. However, during the G₁ stage, the cell is quite active at the biochemical level. The cell is accumulating the building blocks of chromosomal DNA and the associated proteins as well as accumulating sufficient energy reserves to complete the task of replicating each chromosome in the nucleus.

S Phase (Synthesis of DNA)

Throughout interphase, nuclear DNA remains in a semi-condensed chromatin configuration. In the **S phase**, DNA replication can proceed through the mechanisms that result in the formation of identical pairs of DNA molecules—sister chromatids—that are firmly attached to the centromeric region. The centrosome is also duplicated during the S phase. The two centrosomes of homologous chromosomes will give rise to the **mitotic spindle**, the apparatus that orchestrates the movement of chromosomes during mitosis. For example, roughly at the center of each animal cell, the centrosomes are associated with a pair of rod-like objects, the **centrioles**, which are positioned at right angles to each other. Centrioles help organize cell division. We should note, however, that centrioles are not present in the centrosomes of other eukaryotic organisms, such as plants and most fungi.

G₂ Phase (Second Gap)

In the **G₂ phase**, the cell replenishes its energy stores and synthesizes proteins necessary for chromosome manipulation and movement. Some cell organelles are duplicated, and the cytoskeleton is dismantled to provide resources for the mitotic phase. There may be additional cell growth during G₂. The final preparations for the mitotic phase must be completed before the cell is able to enter the first stage of mitosis.

During prometaphase, mitotic spindle microtubules from opposite poles attach to each sister chromatid at the kinetochore. In anaphase, the connection between the sister chromatids breaks down, and the microtubules pull the chromosomes toward opposite poles. During cytokinesis in animal cells, a ring of actin filaments forms at the metaphase plate. The ring contracts, forming a cleavage furrow, which divides the cell in two. In plant cells, Golgi vesicles coalesce at the former metaphase plate, forming a phragmoplast. A cell plate formed by the fusion of the vesicles of the phragmoplast grows from the center toward the cell walls, and the membranes of the vesicles fuse to form a plasma membrane that divides the cell in two.

The Mitotic Phase

The mitotic phase is a multistep process during which the duplicated chromosomes are aligned,

separated, and move into two new, identical daughter cells. The first portion of the mitotic phase is called **karyokinesis**, or nuclear division. As we have just seen, the second portion of the mitotic phase (and often viewed as a process separate from and following mitosis) is called cytokinesis—the physical separation of the cytoplasmic components into the two daughter cells.

Link to Learning

Revisit the stages of mitosis at this [site](#).

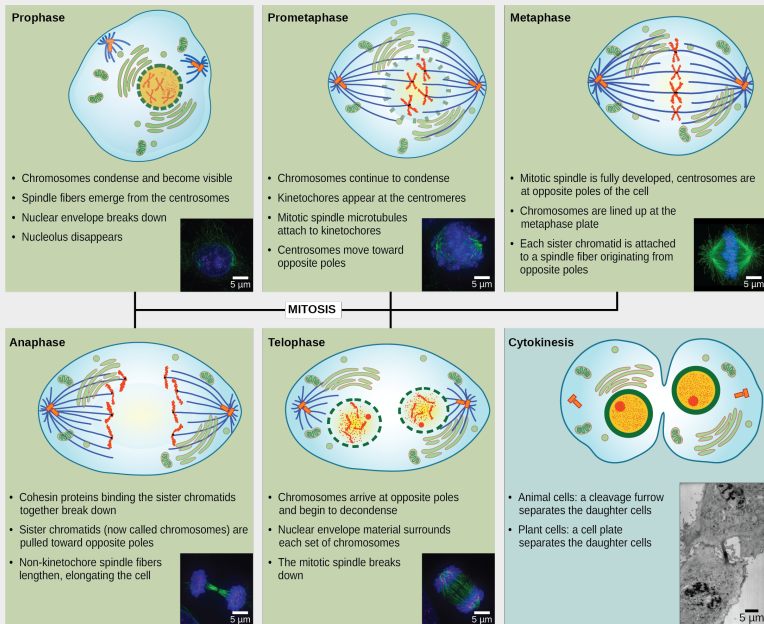
Karyokinesis (Mitosis)

Karyokinesis, also known as **mitosis**, is divided into a series of phases—prophase, prometaphase, metaphase, anaphase, and telophase—that result in the division of the cell nucleus ([link](#)).

Visual Connection

Karyokinesis (or mitosis) is divided into five stages—prophase, prometaphase, metaphase, anaphase, and telophase. The pictures at the bottom were taken by fluorescence microscopy (hence, the black background) of cells artificially stained by

fluorescent dyes: blue fluorescence indicates DNA (chromosomes) and green fluorescence indicates microtubules (spindle apparatus). (credit “mitosis drawings”: modification of work by Mariana Ruiz Villareal; credit “micrographs”: modification of work by Roy van Heesbeen; credit “cytokinesis micrograph”: Wadsworth Center/New York State Department of Health; scale-bar data from Matt Russell)



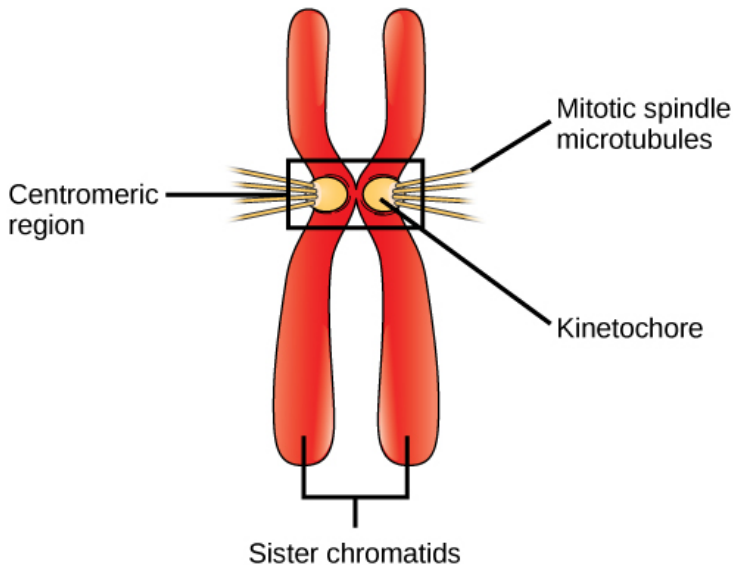
Which of the following is the correct order of events in mitosis?

1. Sister chromatids line up at the metaphase plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.

2. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
3. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
4. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

Prophase (the “first phase”): the nuclear envelope starts to dissociate into small vesicles, and the membranous organelles (such as the Golgi complex [Golgi apparatus] and the endoplasmic reticulum), fragment and disperse toward the periphery of the cell. The nucleolus disappears (dispersed) as well, and the centrosomes begin to move to opposite poles of the cell. Microtubules that will form the *mitotic spindle* extend between the centrosomes, *pushing them farther apart* as the microtubule fibers lengthen. The sister chromatids begin to coil more tightly with the aid of **condensin proteins** and now become visible under a light microscope.

Prometaphase (the “first change phase”): Many processes that began in prophase continue to advance. The remnants of the nuclear envelope fragment further, and the mitotic spindle continues to develop as more microtubules assemble and stretch across the length of the former nuclear area. Chromosomes become even more condensed and discrete. Each sister chromatid develops a protein structure called a **kinetochore** in its centromeric region ([\[link\]](#)). The proteins of the kinetochore attract and bind to the mitotic spindle microtubules. As the spindle microtubules extend from the centrosomes, some of these microtubules come into contact with and firmly bind to the kinetochores. Once a mitotic fiber attaches to a chromosome, the chromosome will be oriented until the kinetochores of sister chromatids face the *opposite poles*. Eventually, all the sister chromatids will be attached via their kinetochores to microtubules from opposing poles. Spindle microtubules that do not engage the chromosomes are called **polar microtubules**. These microtubules overlap each other midway between the two poles and contribute to *cell elongation*. Astral microtubules are located near the poles, aid in spindle orientation, and are required for the regulation of mitosis.



Metaphase (the “change phase”): All the chromosomes are aligned in a plane called the **metaphase plate**, or the equatorial plane, roughly midway between the two poles of the cell. The sister chromatids are still tightly attached to each other by cohesin proteins. At this time, the chromosomes are maximally condensed.

Anaphase (“upward phase”): The cohesin proteins degrade, and the sister chromatids separate at the centromere. Each chromatid, now called a single chromosome, is pulled rapidly toward the centrosome to which its microtubule is attached. The cell becomes visibly elongated (oval shaped) as the polar microtubules slide against each other at the metaphase plate where they overlap.

Telophase (the “distance phase”): the chromosomes reach the opposite poles and begin to *decondense* (unravel), relaxing once again into a stretched-out chromatin configuration. The mitotic spindles are depolymerized into tubulin monomers that will be used to assemble cytoskeletal components for each daughter cell. Nuclear envelopes form around the chromosomes, and nucleosomes appear within the nuclear area.

Cytokinesis

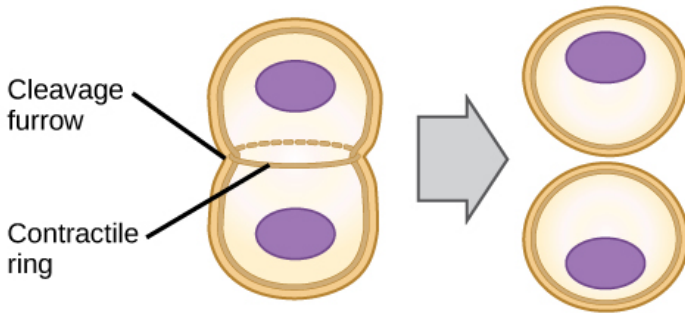
Cytokinesis, or “cell motion,” is sometimes viewed as the second main stage of the mitotic phase, during which cell division is completed via the physical separation of the cytoplasmic components into two daughter cells. However, as we have seen earlier, cytokinesis can also be viewed as a separate phase, which may or may not take place following mitosis. If cytokinesis does take place, cell division is not complete until the cell components have been apportioned and completely separated into the two daughter cells. Although the stages of mitosis are similar for most eukaryotes, the process of cytokinesis is quite different for eukaryotes that have cell walls, such as plant cells.

In animal cells, cytokinesis typically starts during late anaphase. A contractile ring composed of actin filaments forms just inside the plasma membrane at the former metaphase plate. The actin filaments pull

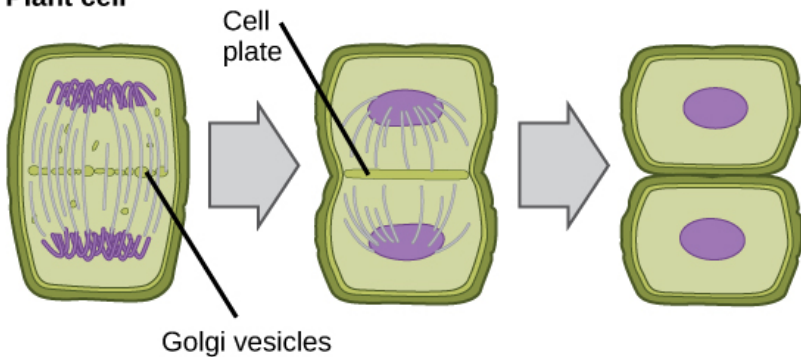
the equator of the cell inward, forming a fissure. This fissure is called the **cleavage furrow**. The furrow deepens as the actin ring contracts, and eventually the membrane is cleaved in two ([\[link\]](#)).

In plant cells, a new cell wall must form between the daughter cells. During interphase, the Golgi apparatus accumulates enzymes, structural proteins, and glucose molecules prior to breaking into vesicles and dispersing throughout the dividing cell. During telophase, these Golgi vesicles are transported on microtubules to form a *phragmoplast* (a vesicular structure) at the metaphase plate. There, the vesicles fuse and coalesce from the center toward the cell walls; this structure is called a **cell plate**. As more vesicles fuse, the cell plate enlarges until it merges with the cell walls at the periphery of the cell. Enzymes use the glucose that has accumulated between the membrane layers to build a new cell wall. The Golgi membranes become parts of the plasma membrane on either side of the new cell wall ([\[link\]](#)).

Animal cell



Plant cell



Go Phase

Not all cells adhere to the classic cell-cycle pattern in which a newly formed daughter cell immediately enters the preparatory phases of interphase, closely followed by the mitotic phase, and cytokinesis. Cells in **Go phase** are not actively preparing to divide. The cell is in a **quiescent** (inactive) stage that occurs when cells exit the cell cycle. Some cells enter Go temporarily due to environmental conditions such as availability of nutrients, or

stimulation by growth factors. The cell will remain in this phase until conditions improve or until an external signal triggers the onset of G₁. Other cells that never or rarely divide, such as mature cardiac muscle and nerve cells, remain in G₀ permanently.

Scientific Method Connection

Determine the Time Spent in Cell-Cycle Stages

Problem: How long does a cell spend in interphase compared to each stage of mitosis?

Background: A prepared microscope slide of whitefish blastula cross-sections will show cells arrested in various stages of the cell cycle. (Note: It is not visually possible to separate the stages of interphase from each other, but the mitotic stages are readily identifiable.) If 100 cells are examined, the number of cells in each identifiable cell-cycle stage will give an estimate of the time it takes for the cell to complete that stage.

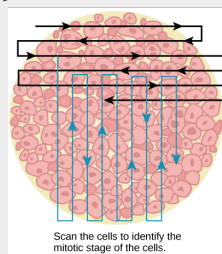
Problem Statement: Given the events included in all of interphase and those that take place in each stage of mitosis, estimate the length of each stage based on a 24-hour cell cycle. Before proceeding, state your hypothesis.

Test your hypothesis: Test your hypothesis by doing the following:

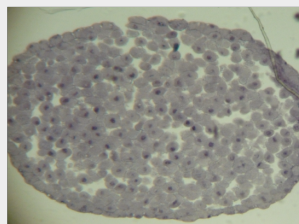
1. Place a fixed and stained microscope slide of whitefish blastula cross-sections under the

scanning objective of a light microscope.

2. Locate and focus on one of the sections using the low-power objective of your microscope. Notice that the section is a circle composed of dozens of closely packed individual cells.
3. Switch to the medium-power objective and refocus. With this objective, individual cells are clearly visible, but the chromosomes will still be very small.
4. Switch to the high-power objective and slowly move the slide left to right, and up and down to view all the cells in the section ([\[link\]](#)). As you scan, you will notice that most of the cells are not undergoing mitosis but are in the interphase period of the cell cycle. Slowly scan whitefish blastula cells with the high-power objective as illustrated in image (a) to identify their mitotic stage. (b) A microscopic image of the scanned cells is shown. (credit “micrograph”: modification of work by Linda Flora; scale-bar data from Matt Russell)



(a)



(b)

5. Practice identifying the various stages of the cell cycle, using the drawings of the stages as

a guide ([\[link\]](#)).

6. Once you are confident about your identification, begin to record the stage of each cell you encounter as you scan left to right, and top to bottom across the blastula section.
7. Keep a tally of your observations and stop when you reach 100 cells identified.
8. The larger the sample size (total number of cells counted), the more accurate the results. If possible, gather and record group data prior to calculating percentages and making estimates.

Record your observations: Make a table similar to [\[link\]](#) within which to record your observations.

Results of Cell Stage

Identification

Phase or Stage	Individual Totals	Group Totals	Percent
Interphase			
Prophase			
Metaphase			
Anaphase			
Telophase			

Cytokinesis						
Totals	100	100	100 percent			

Analyze your data/report your results: To find the length of time whitefish blastula cells spend in each stage, multiply the percent (recorded as a decimal) by 24 hours. Make a table similar to [\[link\]](#) to illustrate your data.

Estimate of Cell

Stage Length

Phase or Stage	Percent	Time in Hours
----------------	---------	---------------

Interphase		
------------	--	--

Prophase		
----------	--	--

Metaphase		
-----------	--	--

Anaphase		
----------	--	--

Telophase		
-----------	--	--

Cytokinesis		
-------------	--	--

Draw a conclusion: Did your results support your estimated times? Were any of the outcomes unexpected? If so, discuss those events in that stage that may have contributed to the calculated time.

Section Summary

The cell cycle is an orderly sequence of events. Cells on the path to cell division proceed through a series of precisely timed and carefully regulated stages. In eukaryotes, the cell cycle consists of a long preparatory period, called interphase, during which chromosomes are replicated. Interphase is divided into G₁, S, and G₂ phases. The mitotic phase begins with karyokinesis (mitosis), which consists of five stages: prophase, prometaphase, metaphase, anaphase, and telophase. The final stage of the cell division process, and sometimes viewed as the final stage of the mitotic phase, is cytokinesis, during which the cytoplasmic components of the daughter cells are separated either by an actin ring (animal cells) or by cell plate formation (plant cells).

Review Questions

Chromosomes are duplicated during what stage of the cell cycle?

1. G₁ phase
2. S phase
3. prophase

4. prometaphase

B

Which of the following events does not occur during some stages of interphase?

1. DNA duplication
 2. organelle duplication
 3. increase in cell size
 4. separation of sister chromatids
-

D

The mitotic spindles arise from which cell structure?

1. centromere
 2. centrosome
 3. kinetochore
 4. cleavage furrow
-

B

Attachment of the mitotic spindle fibers to the kinetochores is a characteristic of which stage

of mitosis?

1. prophase
 2. prometaphase
 3. metaphase
 4. anaphase
-

B

Unpacking of chromosomes and the formation of a new nuclear envelope is a characteristic of which stage of mitosis?

1. prometaphase
 2. metaphase
 3. anaphase
 4. telophase
-

D

Separation of the sister chromatids is a characteristic of which stage of mitosis?

1. prometaphase
 2. metaphase
 3. anaphase
 4. telophase
-

C

The chromosomes become visible under a light microscope during which stage of mitosis?

1. prophase
2. prometaphase
3. metaphase
4. anaphase

A

The fusing of Golgi vesicles at the metaphase plate of dividing plant cells forms what structure?

1. cell plate
2. actin ring
3. cleavage furrow
4. mitotic spindle

A

[\[link\]](#) Which of the following is the correct order of events in mitosis?

1. Sister chromatids line up at the metaphase

plate. The kinetochore becomes attached to the mitotic spindle. The nucleus reforms and the cell divides. Cohesin proteins break down and the sister chromatids separate.

2. The kinetochore becomes attached to the mitotic spindle. Cohesin proteins break down and the sister chromatids separate. Sister chromatids line up at the metaphase plate. The nucleus reforms and the cell divides.
3. The kinetochore becomes attached to the cohesin proteins. Sister chromatids line up at the metaphase plate. The kinetochore breaks down and the sister chromatids separate. The nucleus reforms and the cell divides.
4. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

[\[link\]](#) D. The kinetochore becomes attached to the mitotic spindle. Sister chromatids line up at the metaphase plate. Cohesin proteins break down and the sister chromatids separate. The nucleus reforms and the cell divides.

Critical Thinking Questions

Briefly describe the events that occur in each phase of interphase.

During G₁, the cell increases in size, the genomic DNA is assessed for damage, and the cell stockpiles energy reserves and the components to synthesize DNA. During the S phase, the chromosomes, the centrosomes, and the centrioles (animal cells) duplicate. During the G₂ phase, the cell recovers from the S phase, continues to grow, duplicates some organelles, and dismantles other organelles.

Chemotherapy drugs such as *vincristine* (derived from Madagascar periwinkle plants) and *colchicine* (derived from autumn crocus plants) disrupt mitosis by binding to tubulin (the subunit of microtubules) and interfering with microtubule assembly and disassembly. Exactly what mitotic structure is targeted by these drugs and what effect would that have on cell division?

The mitotic spindle is formed of microtubules. Microtubules are polymers of the protein

tubulin; therefore, it is the mitotic spindle that is disrupted by these drugs. Without a functional mitotic spindle, the chromosomes will not be sorted or separated during mitosis. The cell will arrest in mitosis and die.

Describe the similarities and differences between the cytokinesis mechanisms found in animal cells versus those in plant cells.

There are very few similarities between animal cell and plant cell cytokinesis. In animal cells, a ring of actin fibers is formed around the periphery of the cell at the former metaphase plate (cleavage furrow). The actin ring contracts inward, pulling the plasma membrane toward the center of the cell until the cell is pinched in two. In plant cells, a new cell wall must be formed between the daughter cells. Due to the rigid cell walls of the parent cell, contraction of the middle of the cell is not possible. Instead, a phragmoplast first forms. Subsequently, a cell plate is formed in the center of the cell at the former metaphase plate. The cell plate is formed from Golgi vesicles that contain enzymes, proteins, and glucose. The vesicles fuse and the enzymes build a new cell wall from the proteins and glucose. The cell plate grows toward and eventually fuses with the cell wall of the parent cell.

List some reasons why a cell that has just completed cytokinesis might enter the G₀ phase instead of the G₁ phase.

Many cells temporarily enter G₀ until they reach maturity. Some cells are only triggered to enter G₁ when the organism needs to increase that particular cell type. Some cells only reproduce following an injury to the tissue. Some cells never divide once they reach maturity.

What cell-cycle events will be affected in a cell that produces mutated (non-functional) cohesin protein?

If cohesin is not functional, chromosomes are not packaged after DNA replication in the S phase of interphase. It is likely that the proteins of the centromeric region, such as the kinetochore, would not form. Even if the mitotic spindle fibers could attach to the chromatids without packing, the chromosomes would not be sorted or separated during mitosis.

Glossary

anaphase

stage of mitosis during which sister chromatids are separated from each other

cell cycle

ordered series of events involving cell growth and cell division that produces two new daughter cells

cell plate

structure formed during plant cell cytokinesis by Golgi vesicles, forming a temporary structure (phragmoplast) and fusing at the metaphase plate; ultimately leads to the formation of cell walls that separate the two daughter cells

centriole

rod-like structure constructed of microtubules at the center of each animal cell centrosome

cleavage furrow

constriction formed by an actin ring during cytokinesis in animal cells that leads to cytoplasmic division

condensin

proteins that help sister chromatids coil during prophase

cytokinesis

division of the cytoplasm following mitosis that forms two daughter cells.

G₀ phase

distinct from the G₁ phase of interphase; a cell in G₀ is not preparing to divide

G₁ phase

(also, first gap) first phase of interphase centered on cell growth during mitosis

G₂ phase

(also, second gap) third phase of interphase during which the cell undergoes final preparations for mitosis

interphase

period of the cell cycle leading up to mitosis; includes G₁, S, and G₂ phases (the interim period between two consecutive cell divisions)

karyokinesis

mitotic nuclear division

kinetochore

protein structure associated with the centromere of each sister chromatid that attracts and binds spindle microtubules during prometaphase

metaphase plate

equatorial plane midway between the two poles of a cell where the chromosomes align during metaphase

metaphase

stage of mitosis during which chromosomes are aligned at the metaphase plate

mitosis

(also, karyokinesis) period of the cell cycle during which the duplicated chromosomes are separated into identical nuclei; includes prophase, prometaphase, metaphase, anaphase, and telophase

mitotic phase

period of the cell cycle during which duplicated chromosomes are distributed into two nuclei and cytoplasmic contents are divided; includes karyokinesis (mitosis) and cytokinesis

mitotic spindle

apparatus composed of microtubules that orchestrates the movement of chromosomes during mitosis

prometaphase

stage of mitosis during which the nuclear membrane breaks down and mitotic spindle fibers attach to kinetochores

prophase

stage of mitosis during which chromosomes condense and the mitotic spindle begins to form

quiescent

refers to a cell that is performing normal cell functions and has not initiated preparations for cell division

S phase

second, or synthesis, stage of interphase during which DNA replication occurs

telophase

stage of mitosis during which chromosomes arrive at opposite poles, decondense, and are surrounded by a new nuclear envelope

Control of the Cell Cycle

By the end of this section, you will be able to do the following:

- Understand how the cell cycle is controlled by mechanisms that are both internal and external to the cell
- Explain how the three internal “control checkpoints” occur at the end of G₁, at the G₂/M transition, and during metaphase
- Describe the molecules that control the cell cycle through positive and negative regulation

The length of the cell cycle is highly variable, even within the cells of a single organism. In humans, the frequency of cell turnover ranges from a few hours in early embryonic development, to an average of two to five days for epithelial cells, and to an entire human lifetime spent in G₀ by specialized cells, such as cortical neurons or cardiac muscle cells.

There is also variation in the time that a cell spends in each phase of the cell cycle. When rapidly dividing mammalian cells are grown in a culture (outside the body under optimal growing conditions), the length of the cell cycle is about 24 hours. In rapidly dividing human cells with a 24-hour cell cycle, the G₁ phase lasts approximately nine hours, the S phase lasts 10 hours, the G₂ phase lasts about four and one-half hours, and the M phase lasts approximately one-half hour. By comparison,

in fertilized eggs (and early embryos) of fruit flies, the cell cycle is completed in about eight minutes. This is because the nucleus of the fertilized egg divides many times by mitosis but does not go through cytokinesis until a multinucleate “zygote” has been produced, with many nuclei located along the periphery of the cell membrane, thereby shortening the time of the cell division cycle. The timing of events in the cell cycle of both “invertebrates” and “vertebrates” is controlled by mechanisms that are both internal and external to the cell.

Regulation of the Cell Cycle by External Events

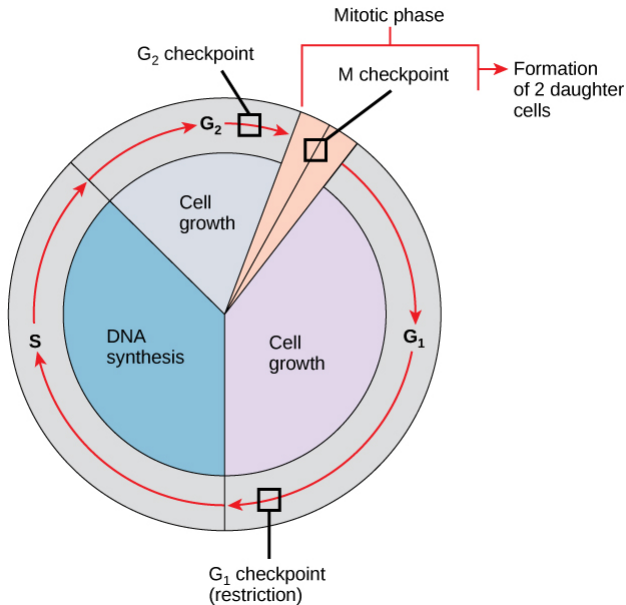
Both the initiation and inhibition of cell division are triggered by events external to the cell when it is about to begin the replication process. An event may be as simple as the death of nearby cells or as sweeping as the release of growth-promoting hormones, such as **human growth hormone (HGH or hGH)**. A lack of HGH can *inhibit* cell division, resulting in dwarfism, whereas too much HGH can result in gigantism. Crowding of cells can also inhibit cell division. In contrast, a factor that can initiate cell division is the size of the cell: As a cell grows, it becomes physiologically inefficient due to its decreasing surface-to-volume ratio. The solution to this problem is to divide.

Whatever the source of the message, the cell receives the signal, and a series of events within the cell allows it to proceed into interphase. Moving forward from this initiation point, every parameter required during each cell cycle phase must be met or the cycle cannot progress.

The cell cycle is controlled at three checkpoints. The integrity of the DNA is assessed at the G₁ checkpoint. Proper chromosome duplication is assessed at the G₂ checkpoint. Attachment of each kinetochore to a spindle fiber is assessed at the M checkpoint.

Regulation at Internal Checkpoints

It is essential that the daughter cells produced be exact duplicates of the parent cell. Mistakes in the duplication or distribution of the chromosomes lead to mutations that may be passed forward to every new cell produced from an abnormal cell. To prevent a compromised cell from continuing to divide, there are internal control mechanisms that operate at three main **cell-cycle checkpoints**: A checkpoint is one of several points in the eukaryotic cell cycle at which the progression of a cell to the next stage in the cycle can be halted until conditions are favorable. These checkpoints occur near the end of G₁, at the G₂/M transition, and during metaphase ([\[link\]](#)).



The G₁ Checkpoint

The **G₁ checkpoint** determines whether all conditions are favorable for cell division to proceed. The G₁ checkpoint, also called the restriction point (in yeast), is a point at which the cell irreversibly commits to the cell division process. External influences, such as growth factors, play a large role in carrying the cell past the G₁ checkpoint. In addition to adequate reserves and cell size, there is a check for genomic DNA damage at the G₁ checkpoint. A cell that does not meet all the requirements will not be allowed to progress into the S phase. The cell can halt the cycle and attempt to remedy the problematic condition, or the cell can advance into G₀ and await further signals when

conditions improve.

The G₂ Checkpoint

The G₂ checkpoint bars entry into the mitotic phase if certain conditions are not met. As at the G₁ checkpoint, cell size and protein reserves are assessed. However, the most important role of the G₂ checkpoint is to ensure that all of the chromosomes have been replicated and that the replicated DNA is not damaged. If the checkpoint mechanisms detect problems with the DNA, the cell cycle is halted, and the cell attempts to either complete DNA replication or repair the damaged DNA.

The M Checkpoint

The M checkpoint occurs near the end of the metaphase stage of karyokinesis. The M checkpoint is also known as the spindle checkpoint, because it determines whether all the sister chromatids are correctly attached to the spindle microtubules. Because the separation of the sister chromatids during anaphase is an irreversible step, the cycle will not proceed until the kinetochores of each pair of sister chromatids are firmly anchored to at least two spindle fibers arising from opposite poles of the cell.

Link to Learning

Watch what occurs at the G₁, G₂, and M checkpoints by visiting this [website](#) to see an animation of the cell cycle.

The concentrations of cyclin proteins change throughout the cell cycle. There is a direct correlation between cyclin accumulation and the three major cell-cycle checkpoints. Also note the sharp decline of cyclin levels following each checkpoint (the transition between phases of the cell cycle), as cyclin is degraded by cytoplasmic enzymes. (credit: modification of work by "WikiMiMa"/Wikimedia Commons) *Cyclin-dependent kinases (Cdks)* are protein kinases that, when fully activated, can phosphorylate and thus activate other proteins that advance the cell cycle past a checkpoint. To become fully activated, a Cdk must bind to a cyclin protein and then be phosphorylated by another kinase.

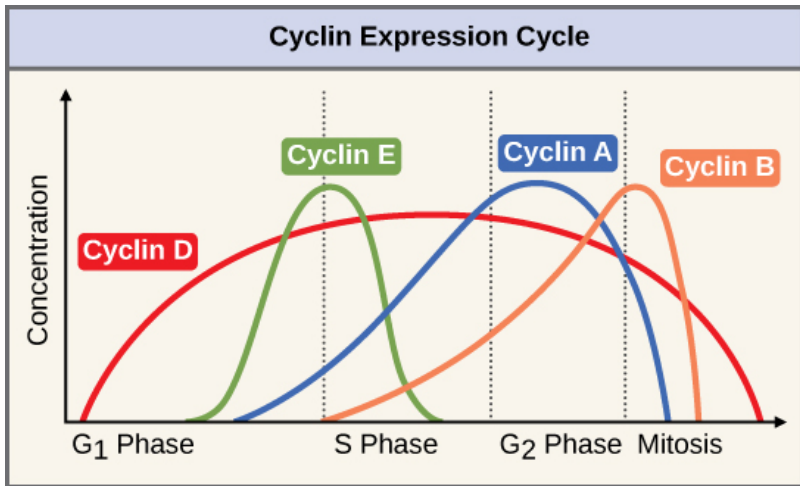
Regulator Molecules of the Cell Cycle

In addition to the internally controlled checkpoints, there are two groups of intracellular molecules that regulate the cell cycle. These regulatory molecules either promote progress of the cell to the next phase (positive regulation) or halt the cycle (negative

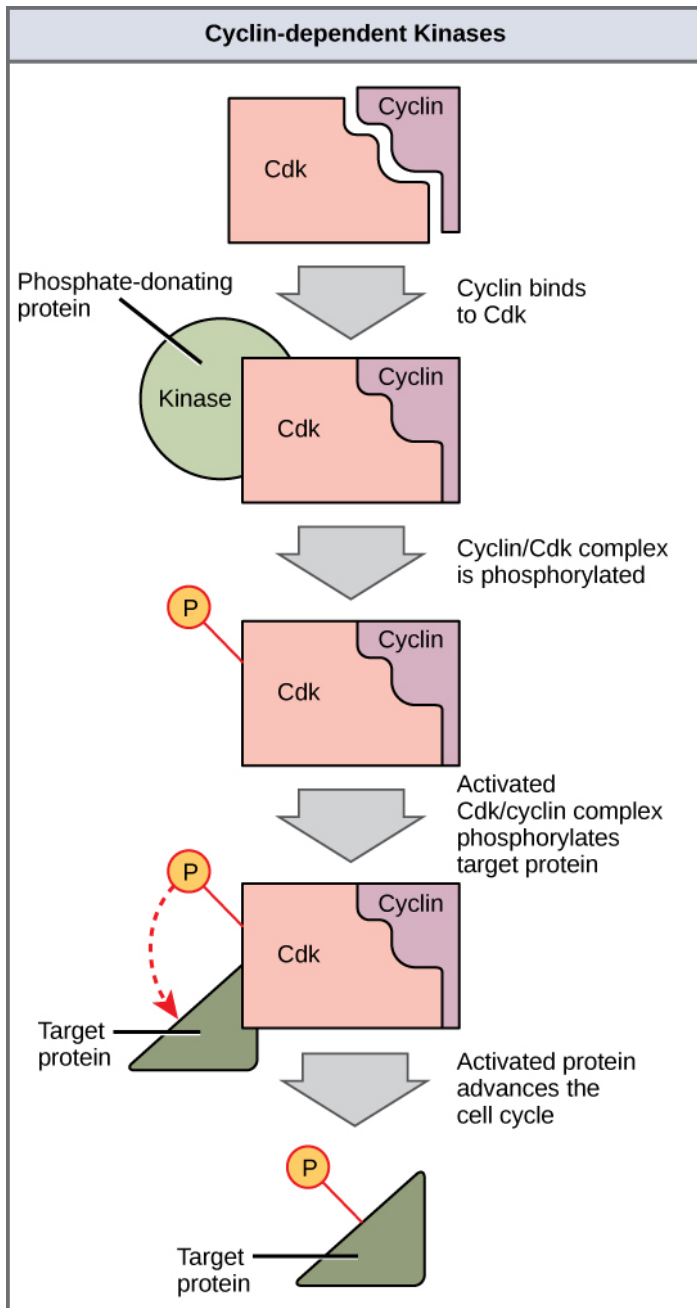
regulation). Regulator molecules may act individually, or they can influence the activity or production of other regulatory proteins. Therefore, the failure of a single regulator may have almost no effect on the cell cycle, especially if more than one mechanism controls the same event. However, the effect of a deficient or non-functioning regulator can be wide-ranging and possibly fatal to the cell if multiple processes are affected.

Positive Regulation of the Cell Cycle

Two groups of proteins, called **cyclins** and **cyclin-dependent kinases** (Cdks), are termed positive regulators. They are responsible for the progress of the cell through the various checkpoints. The levels of the four cyclin proteins fluctuate throughout the cell cycle in a predictable pattern ([\[link\]](#)). Increases in the concentration of cyclin proteins are triggered by both external and internal signals. After the cell moves to the next stage of the cell cycle, the cyclins that were active in the previous stage are degraded by cytoplasmic enzymes, as shown in [\[link\]](#) below.



Cyclins regulate the cell cycle only when they are tightly bound to Cdks. To be fully active, the Cdk/cyclin complex must also be phosphorylated in specific locations to activate the complex. Like all kinases, Cdks are enzymes (*kinases*) that in turn phosphorylate other proteins. Phosphorylation activates the protein by changing its shape. The proteins phosphorylated by Cdks are involved in advancing the cell to the next phase. ([link]). The levels of Cdk proteins are relatively stable throughout the cell cycle; however, the concentrations of cyclin fluctuate and determine when Cdk/cyclin complexes form. The different cyclins and Cdks bind at specific points in the cell cycle and thus regulate different checkpoints.



Because the cyclic fluctuations of cyclin levels are

largely based on the *timing of the cell cycle* and not on specific events, regulation of the cell cycle usually occurs by either the Cdk molecules alone or the Cdk/cyclin complexes. Without a specific concentration of fully activated cyclin/Cdk complexes, the cell cycle cannot proceed through the checkpoints.

Although the cyclins are the main regulatory molecules that determine the forward momentum of the cell cycle, there are several other mechanisms that fine-tune the progress of the cycle with negative, rather than positive, effects. These mechanisms essentially block the progression of the cell cycle until problematic conditions are resolved. Molecules that prevent the full activation of Cdks are called Cdk inhibitors. Many of these inhibitor molecules directly or indirectly monitor a particular cell-cycle event. The block placed on Cdks by inhibitor molecules will not be removed until the specific event that the inhibitor monitors is completed.

Negative Regulation of the Cell Cycle

The second group of cell-cycle regulatory molecules are *negative regulators*, which stop the cell cycle. Remember that in positive regulation, active molecules cause the cycle to progress.

The best understood negative regulatory molecules

are **retinoblastoma protein (Rb)**, **p53**, and **p21**. Retinoblastoma proteins are a group of *tumor-suppressor proteins* common in many cells. We should note here that the 53 and 21 designations refer to the functional molecular masses of the proteins (p) in kilodaltons (a dalton is equal to an *atomic mass unit*, which is equal to one proton or one neutron or 1 g/mol). Much of what is known about cell-cycle regulation comes from research conducted with cells that have *lost regulatory control*. All three of these regulatory proteins were discovered to be damaged or non-functional in cells that had begun to replicate uncontrollably (i.e., became cancerous). In each case, the main cause of the unchecked progress through the cell cycle was a faulty copy of the regulatory protein.

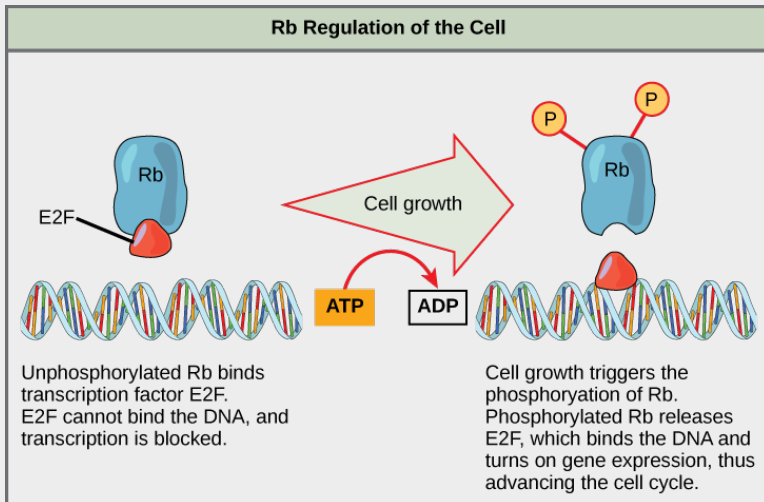
Rb, p53, and p21 act primarily at the G₁ checkpoint. p53 is a multi-functional protein that has a major impact on the commitment of a cell to division because it acts when there is damaged DNA in cells that are undergoing the preparatory processes during G₁. If damaged DNA is detected, p53 halts the cell cycle and then recruits specific enzymes to repair the DNA. If the DNA cannot be repaired, p53 can trigger apoptosis, or cell suicide, to prevent the duplication of damaged chromosomes. As p53 levels rise, the production of p21 is triggered. p21 enforces the halt in the cycle dictated by p53 by binding to and inhibiting the activity of the Cdk/cyclin complexes. As a cell is exposed to more stress,

higher levels of p53 and p21 accumulate, making it less likely that the cell will move into the S phase.

Rb, which largely monitors cell size, exerts its regulatory influence on other positive regulator proteins. In the *active*, dephosphorylated state, Rb binds to proteins called *transcription factors*, most commonly, E2F ([\[link\]](#)). Transcription factors “turn on” specific genes, allowing the production of proteins encoded by that gene. When Rb is bound to E2F, production of proteins necessary for the G₁/S transition is blocked. As the cell increases in size, Rb is slowly phosphorylated until it becomes *inactivated*. Rb releases E2F, which can now turn on the gene that produces the transition protein, and this particular block is removed. For the cell to move past each of the checkpoints, all positive regulators must be “turned on,” and all negative regulators must be “turned off.”

Visual Connection

Rb halts the cell cycle and releases its hold in response to cell growth.



Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

Section Summary

Each step of the cell cycle is monitored by internal controls called checkpoints. There are three major checkpoints in the cell cycle: one near the end of G₁, a second at the G₂/M transition, and the third during metaphase. Positive regulator molecules allow the cell cycle to advance to the next stage of cell division. Negative regulator molecules monitor cellular conditions and can halt the cycle until specific requirements are met.

Visual Connection Questions

[\[link\]](#) Rb and other proteins that negatively regulate the cell cycle are sometimes called tumor suppressors. Why do you think the name tumor suppressor might be appropriate for these proteins?

[\[link\]](#) Rb and other negative regulatory proteins control cell division and therefore prevent the formation of tumors. Mutations that prevent these proteins from carrying out their function can result in cancer.

Review Questions

At which of the cell-cycle checkpoints do external forces have the greatest influence?

1. G₁ checkpoint
2. G₂ checkpoint
3. M checkpoint
4. G₀ checkpoint

A

What is the main prerequisite for clearance at the G₂ checkpoint?

1. cell has reached a sufficient size
2. an adequate stockpile of nucleotides
3. accurate and complete DNA replication
4. proper attachment of mitotic spindle fibers to kinetochores

C

If the M checkpoint is not cleared, what stage of mitosis will be blocked?

1. prophase
2. prometaphase
3. metaphase
4. anaphase

D

Which protein is a positive regulator that phosphorylates other proteins when activated?

1. p53

2. retinoblastoma protein (Rb)
 3. cyclin
 4. cyclin-dependent kinase (Cdk)
-

D

Many of the negative regulator proteins of the cell cycle were discovered in what type of cells?

1. gametes
 2. cells in G₀
 3. cancer cells
 4. stem cells
-

C

Which negative regulatory molecule can trigger cell suicide (apoptosis) if vital cell cycle events do not occur?

1. p53
 2. p21
 3. retinoblastoma protein (Rb)
 4. cyclin-dependent kinase (Cdk)
-

A

Critical Thinking Questions

Describe the general conditions that must be met at each of the three main cell-cycle checkpoints.

The G₁ checkpoint monitors adequate cell growth, the state of the genomic DNA, adequate stores of energy, and materials for S phase. At the G₂ checkpoint, DNA is checked to ensure that all chromosomes were duplicated and that there are no mistakes in newly synthesized DNA. Additionally, cell size and energy reserves are evaluated. The M checkpoint confirms the correct attachment of the mitotic spindle fibers to the kinetochores.

Compare and contrast the roles of the positive cell-cycle regulators negative regulators.

Positive cell regulators such as cyclin and Cdk perform tasks that advance the cell cycle to the next stage. Negative regulators such as Rb, p53, and p21 block the progression of the cell cycle until certain events have occurred.

What steps are necessary for Cdk to become fully active?

Cdk must bind to a cyclin, and it must be phosphorylated in the correct position to become fully active.

Rb is a negative regulator that blocks the cell cycle at the G₁ checkpoint until the cell achieves a requisite size. What molecular mechanism does Rb employ to halt the cell cycle?

Rb is active when it is dephosphorylated. In this state, Rb binds to E2F, which is a transcription factor required for the transcription and eventual translation of molecules required for the G₁/S transition. E2F cannot transcribe certain genes when it is bound to Rb. As the cell increases in size, Rb becomes phosphorylated, inactivated, and releases E2F. E2F can then promote the transcription of the genes it controls, and the transition proteins will be produced.

Glossary

cell-cycle checkpoint

mechanism that monitors the preparedness of a eukaryotic cell to advance through the various cell-cycle stages

cyclin

one of a group of proteins that act in conjunction with cyclin-dependent kinases to help regulate the cell cycle by phosphorylating key proteins; the concentrations of cyclins fluctuate throughout the cell cycle

cyclin-dependent kinase (Cdk)

one of a group of protein kinases that helps to regulate the cell cycle when bound to cyclin; it functions to phosphorylate other proteins that are either activated or inactivated by phosphorylation

p21

cell-cycle regulatory protein that inhibits the cell cycle; its levels are controlled by p53

p53

cell-cycle regulatory protein that regulates cell growth and monitors DNA damage; it halts the progression of the cell cycle in cases of DNA damage and may induce apoptosis

retinoblastoma protein (Rb)

regulatory molecule that exhibits negative effects on the cell cycle by interacting with a

transcription factor (E2F)

Cancer and the Cell Cycle

By the end of this section, you will be able to do the following:

- Describe how cancer is caused by uncontrolled cell growth
- Understand how proto-oncogenes are normal cell genes that, when mutated, become oncogenes
- Describe how tumor suppressors function
- Explain how mutant tumor suppressors cause cancer

Cancer comprises many different diseases caused by a common mechanism: uncontrolled cell growth. Despite the redundancy and overlapping levels of cell-cycle control, errors do occur. One of the critical processes monitored by the cell-cycle checkpoint surveillance mechanism is the proper replication of DNA during the S phase. Even when all of the cell-cycle controls are fully functional, a small percentage of replication errors (mutations) will be passed on to the daughter cells. If changes to the DNA nucleotide sequence occur within a coding portion of a gene and are not corrected, a gene mutation results. All cancers start when a gene mutation gives rise to a faulty protein that plays a key role in cell reproduction.

The change in the cell that results from the malformed protein may be minor: perhaps a slight

delay in the binding of Cdk to cyclin or an Rb protein that detaches from its target DNA while still phosphorylated. Even minor mistakes, however, may allow subsequent mistakes to occur more readily. Over and over, small uncorrected errors are passed from the parent cell to the daughter cells and amplified as each generation produces more non-functional proteins from uncorrected DNA damage. Eventually, the pace of the cell cycle speeds up as the effectiveness of the control and repair mechanisms decreases. Uncontrolled growth of the mutated cells outpaces the growth of normal cells in the area, and a tumor (“-oma”) can result.

Proto-oncogenes

The genes that code for the positive cell-cycle regulators are called **proto-oncogenes**. Proto-oncogenes are normal genes that, when mutated in certain ways, become **oncogenes**—genes that cause a cell to become cancerous. Consider what might happen to the cell cycle in a cell with a recently acquired oncogene. In most instances, the alteration of the DNA sequence will result in a less functional (or non-functional) protein. The result is detrimental to the cell and will likely prevent the cell from completing the cell cycle; however, the organism is not harmed because the mutation will not be carried forward. If a cell cannot reproduce, the mutation is not propagated and the damage is minimal.

Occasionally, however, a gene mutation causes a change that increases the activity of a positive regulator. For example, a mutation that allows Cdk to be activated without being partnered with cyclin could push the cell cycle past a checkpoint before all of the required conditions are met. If the resulting daughter cells are too damaged to undergo further cell divisions, the mutation would not be propagated and no harm would come to the organism. However, if the atypical daughter cells are able to undergo further cell divisions, subsequent generations of cells may accumulate even more mutations, some possibly in additional genes that regulate the cell cycle.

The Cdk gene in the above example is only one of many genes that are considered proto-oncogenes. In addition to the cell-cycle regulatory proteins, any protein that influences the cycle can be altered in such a way as to override cell-cycle checkpoints. An oncogene is any gene that, when altered, leads to an increase in the rate of cell-cycle progression.

Tumor Suppressor Genes

Like proto-oncogenes, many of the negative cell-cycle regulatory proteins were discovered in cells that had become cancerous. **Tumor suppressor genes** are segments of DNA that code for negative regulator proteins, the type of regulators that, when

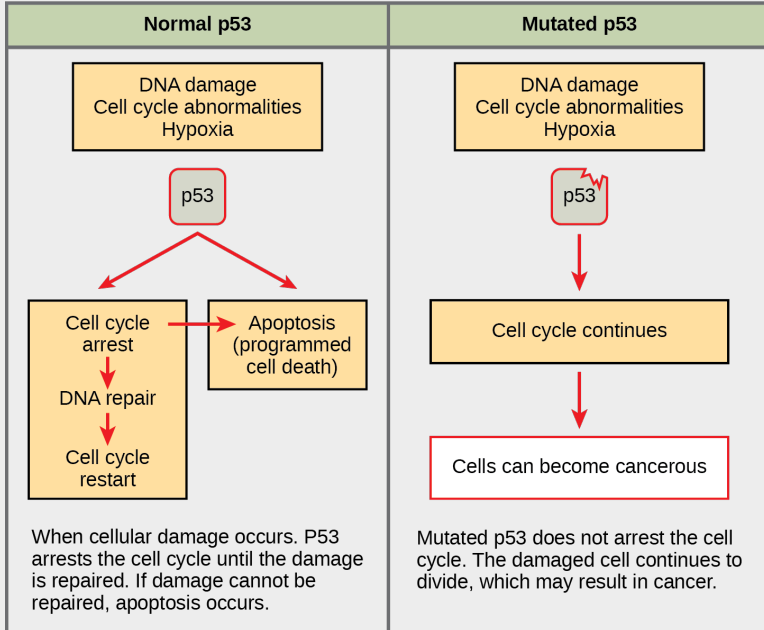
activated, can prevent the cell from undergoing uncontrolled division. The collective function of the best-understood tumor suppressor gene proteins, Rb, p53, and p21, is to put up a roadblock to cell-cycle progression until certain events are completed. A cell that carries a mutated form of a negative regulator might not be able to halt the cell cycle if there is a problem. Tumor suppressors are similar to brakes in a vehicle: Malfunctioning brakes can contribute to a car crash!

Mutated p53 genes have been identified in more than 50 percent of all human tumor cells. This discovery is not surprising in light of the multiple roles that the p53 protein plays at the G₁ checkpoint. A cell with a faulty p53 may fail to detect errors present in the genomic DNA ([\[link\]](#)). Even if a partially functional p53 does identify the mutations, it may no longer be able to signal the necessary DNA repair enzymes. Either way, damaged DNA will remain uncorrected. At this point, a functional p53 will deem the cell unsalvageable and trigger programmed cell death (apoptosis). The damaged version of p53 found in cancer cells, however, cannot trigger apoptosis.

Visual Connection

The role of normal p53 is to monitor DNA and the supply of oxygen (hypoxia is a condition of

reduced oxygen supply). If damage is detected, p53 triggers repair mechanisms. If repairs are unsuccessful, p53 signals apoptosis. A cell with an abnormal p53 protein cannot repair damaged DNA and thus cannot signal apoptosis. Cells with abnormal p53 can become cancerous. (credit: modification of work by Thierry Soussi)



Human papillomavirus can cause cervical cancer. The virus encodes E6, a protein that binds p53. Based on this fact and what you know about p53, what effect do you think E6 binding has on p53 activity?

1. E6 activates p53
2. E6 inactivates p53
3. E6 mutates p53
4. E6 binding marks p53 for degradation

The loss of p53 function has other repercussions for the cell cycle. Mutated p53 might lose its ability to trigger p21 production. Without adequate levels of p21, there is no effective block on Cdk activation. Essentially, without a fully functional p53, the G₁ checkpoint is severely compromised and the cell proceeds directly from G₁ to S regardless of internal and external conditions. At the completion of this shortened cell cycle, two daughter cells are produced that have inherited the mutated p53 gene. Given the non-optimal conditions under which the parent cell reproduced, it is likely that the daughter cells will have acquired other mutations in addition to the faulty tumor-suppressor gene. Cells such as these daughter cells quickly accumulate both oncogenes and non-functional tumor-suppressor genes. Again, the result is tumor growth.

Link to Learning

Watch an animation of how cancer results from errors in the cell cycle.

<https://www.openstax.org/1/cancer>

Section Summary

Cancer is the result of unchecked cell division caused by a breakdown of the mechanisms that regulate the cell cycle. The loss of control begins with a change in the DNA sequence of a gene that codes for one of the regulatory molecules. Faulty instructions lead to a protein that does not function as it should. Any disruption of the monitoring system can allow other mistakes to be passed on to the daughter cells. Each successive cell division will give rise to daughter cells with even more accumulated damage. Eventually, all checkpoints become nonfunctional, and rapidly reproducing cells crowd out normal cells, resulting in a tumor or leukemia (blood cancer).

Visual Connection Questions

[\[link\]](#) Human papillomavirus can cause cervical cancer. The virus encodes E6, a protein that binds p53. Based on this fact and what you know about p53, what effect do you think E6 binding has on p53 activity?

1. E6 activates p53
 2. E6 inactivates p53
 3. E6 mutates p53
 4. E6 binding marks p53 for degradation
-

[\[link\]](#) D. E6 binding marks p53 for degradation.

Review Questions

_____ are changes to the order of nucleotides in a segment of DNA that codes for a protein.

1. Proto-oncogenes
2. Tumor suppressor genes
3. Gene mutations
4. Negative regulators

C

A gene that codes for a positive cell-cycle regulator is called a(n) ____.

1. kinase inhibitor.
2. tumor suppressor gene.
3. proto-oncogene.
4. oncogene.

C

A mutated gene that codes for an altered

version of Cdk that is active in the absence of cyclin is a(n) ____.

1. kinase inhibitor.
2. tumor suppressor gene.
3. proto-oncogene.
4. oncogene.

D

Which molecule is a Cdk inhibitor that is controlled by p53?

1. cyclin
2. anti-kinase
3. Rb
4. p21

D

Critical Thinking Questions

Outline the steps that lead to a cell becoming cancerous.

If one of the genes that produces regulator proteins becomes mutated, it produces a malformed, possibly non-functional, cell-cycle regulator, increasing the chance that more mutations will be left unrepaired in the cell. Each subsequent generation of cells sustains more damage. The cell cycle can speed up as a result of the loss of functional checkpoint proteins. The cells can lose the ability to self-destruct and eventually become “immortalized.”

Explain the difference between a proto-oncogene and a tumor-suppressor gene.

A proto-oncogene is a segment of DNA that codes for one of the positive cell cycle regulators. If that gene becomes mutated so that it produces a hyperactivated protein product, it is considered an oncogene. A tumor suppressor gene is a segment of DNA that codes for one of the negative cell cycle regulators. If that gene becomes mutated so that the protein product becomes less active, the cell cycle will run unchecked. A single oncogene can initiate abnormal cell divisions; however, tumor suppressors lose their effectiveness only when both copies of the gene are damaged.

List the regulatory mechanisms that might be lost in a cell producing faulty p53.

Regulatory mechanisms that might be lost include monitoring of the quality of the genomic DNA, recruiting of repair enzymes, and the triggering of apoptosis.

p53 can trigger apoptosis if certain cell-cycle events fail. How does this regulatory outcome benefit a multicellular organism?

If a cell has damaged DNA, the likelihood of producing faulty proteins is higher. The daughter cells of such a damaged parent cell would also produce faulty proteins that might eventually become cancerous. If p53 recognizes this damage and triggers the cell to self-destruct, the damaged DNA is degraded and recycled. No further harm comes to the organism. Another healthy cell is triggered to divide instead.

Glossary

oncogene

mutated version of a normal gene involved in the positive regulation of the cell cycle

proto-oncogene

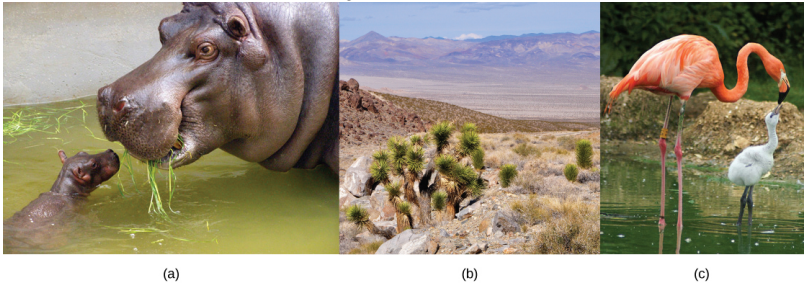
normal gene that when mutated becomes an oncogene

tumor suppressor gene

segment of DNA that codes for regulator proteins that prevent the cell from undergoing uncontrolled division

Introduction

class = "introduction" Each of us, like the organisms shown above, begins life as a fertilized egg (zygote). After trillions of cell divisions, each of us develops into a complex, multicellular organism. (credit a: modification of work by Frank Wouters; credit b: modification of work by Ken Cole, USGS; credit c: modification of work by Martin Pettitt)



The ability to reproduce is a basic characteristic of all organisms: Hippopotamuses give birth to hippopotamus calves; Joshua trees produce seeds from which Joshua tree seedlings emerge; and adult flamingos lay eggs that hatch into flamingo chicks. However, unlike the organisms shown above, offspring may or may not resemble their parents. For example, in the case of most insects such as butterflies (with a complete metamorphosis), the larval forms rarely resemble the adult forms.

Although many unicellular organisms and a few multicellular organisms can produce genetically identical clones of themselves through ***asexual reproduction***, many single-celled organisms and

most multicellular organisms reproduce regularly using another method—***sexual reproduction***. This highly evolved method involves the production by parents of two haploid cells and the fusion of two haploid cells to form a single, genetically recombined diploid cell—a genetically unique organism. Haploid cells that are part of the sexual reproductive cycle are produced by a type of cell division called **meiosis**. Sexual reproduction, involving both meiosis and fertilization, introduces variation into offspring that may account for the evolutionary success of sexual reproduction. The vast majority of eukaryotic organisms, both multicellular and unicellular, can or must employ some form of meiosis and fertilization to reproduce.

In most plants and animals, through thousands of rounds of mitotic cell division, diploid cells (whether produced by asexual or sexual reproduction) will develop into an adult organism.

The Process of Meiosis

By the end of this section, you will be able to do the following:

- Describe the behavior of chromosomes during meiosis, and the differences between the first and second meiotic divisions
- Describe the cellular events that take place during meiosis
- Explain the differences between meiosis and mitosis
- Explain the mechanisms within the meiotic process that produce genetic variation among the haploid gametes

Sexual reproduction requires the union of two specialized cells, called **gametes**, each of which contains one set of chromosomes. When gametes unite, they form a **zygote**, or fertilized egg that contains two sets of chromosomes. (Note: Cells that contain one set of chromosomes are called **haploid**; cells containing two sets of chromosomes are called **diploid**.) If the reproductive cycle is to continue for any sexually reproducing species, then the diploid cell must somehow reduce its number of chromosome sets to produce haploid gametes; otherwise, the number of chromosome sets will double with every future round of fertilization. Therefore, sexual reproduction requires a nuclear division that reduces the number of chromosome sets by half.

Most animals and plants and many unicellular organisms are diploid and therefore have two sets of chromosomes. In each **somatic cell** of the organism (all cells of a multicellular organism except the gametes or reproductive cells), the nucleus contains two copies of each chromosome, called **homologous chromosomes**. Homologous chromosomes are matched pairs containing the same genes in identical locations along their lengths. Diploid organisms inherit one copy of each homologous chromosome from each parent.

Meiosis is the *nuclear division* that forms haploid cells from diploid cells, and it employs many of the same cellular mechanisms as mitosis. However, as you have learned, **mitosis** produces daughter cells whose nuclei are genetically identical to the original parent nucleus. In mitosis, both the parent and the daughter nuclei are at the same “ploidy level”—diploid in the case of most multicellular most animals. Plants use mitosis to grow as sporophytes, and to grow and produce eggs and sperm as gametophytes; so they use mitosis for both haploid and diploid cells (as well as for all other ploidies). In meiosis, the starting nucleus is always diploid and the daughter nuclei that result are haploid. To achieve this reduction in chromosome number, meiosis consists of one round of chromosome replication followed by two rounds of nuclear division. Because the events that occur during each of the division stages are analogous to the events of

mitosis, the same stage names are assigned. However, because there are two rounds of division, the major process and the stages are designated with a “I” or a “II.” Thus, **meiosis I** is the first round of meiotic division and consists of prophase I, prometaphase I, and so on. Likewise, **Meiosis II** (during which the second round of meiotic division takes place) includes prophase II, prometaphase II, and so on.

Early in prophase I, homologous chromosomes come together to form a synapse. The chromosomes are bound tightly together and in perfect alignment by a protein lattice called a synaptonemal complex and by cohesin proteins at the centromere.

Crossover occurs between *nonsister chromatids of homologous chromosomes*. The result is an exchange of genetic material between homologous chromosomes. Random, independent assortment during metaphase I can be demonstrated by considering a cell with a set of two chromosomes ($n = 2$). In this case, there are two possible arrangements at the equatorial plane in metaphase I. The total possible number of different gametes is 2^n , where n equals the number of chromosomes in a set. In this example, there are four possible genetic combinations for the gametes. With $n = 23$ in human cells, there are over eight million possible combinations of paternal and maternal chromosomes.

Meiosis I

Meiosis is preceded by an interphase consisting of G₁, S, and G₂ phases, which are nearly identical to the phases preceding mitosis. The G₁ phase (the “first gap phase”) is focused on cell growth. During the S phase—the second phase of interphase—the cell copies or *replicates* the DNA of the chromosomes. Finally, in the G₂ phase (the “second gap phase”) the cell undergoes the final preparations for meiosis.

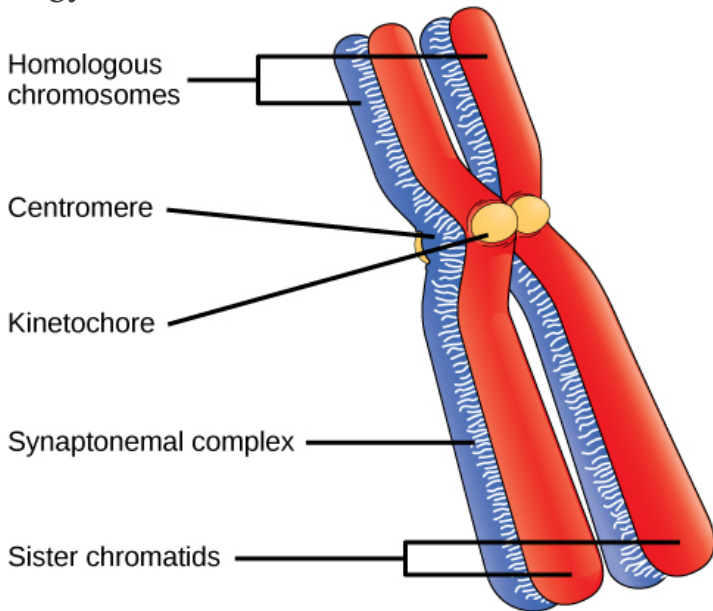
During DNA duplication in the S phase, each chromosome is replicated to produce two identical copies—*sister chromatids* that are held together at the centromere by **cohesin** proteins, which hold the chromatids together until anaphase II.

Prophase I

Early in prophase I, before the chromosomes can be seen clearly with a microscope, the homologous chromosomes are attached at their tips to the nuclear envelope by proteins. As the nuclear envelope begins to break down, the proteins associated with homologous chromosomes bring the pair closer together. Recall that in mitosis, homologous chromosomes do not pair together. The **synaptonemal complex**, a lattice of proteins between the homologous chromosomes, first forms at specific locations and then spreads outward to cover the entire length of the chromosomes. The tight pairing of the homologous chromosomes is

called *synapsis*. In **synapsis**, the genes on the chromatids of the homologous chromosomes are aligned precisely with each other. The synaptonemal complex supports the exchange of chromosomal segments between homologous nonsister chromatids—a process called **crossing over**. Crossing over can be observed visually after the exchange as *chiasmata* (singular = chiasma) ([link](#)).

In humans, even though the X and Y sex chromosomes are not completely homologous (that is, most of their genes differ), there is a small region of homology that allows the X and Y chromosomes to pair up during prophase I. A partial synaptonemal complex develops only between the regions of homology.

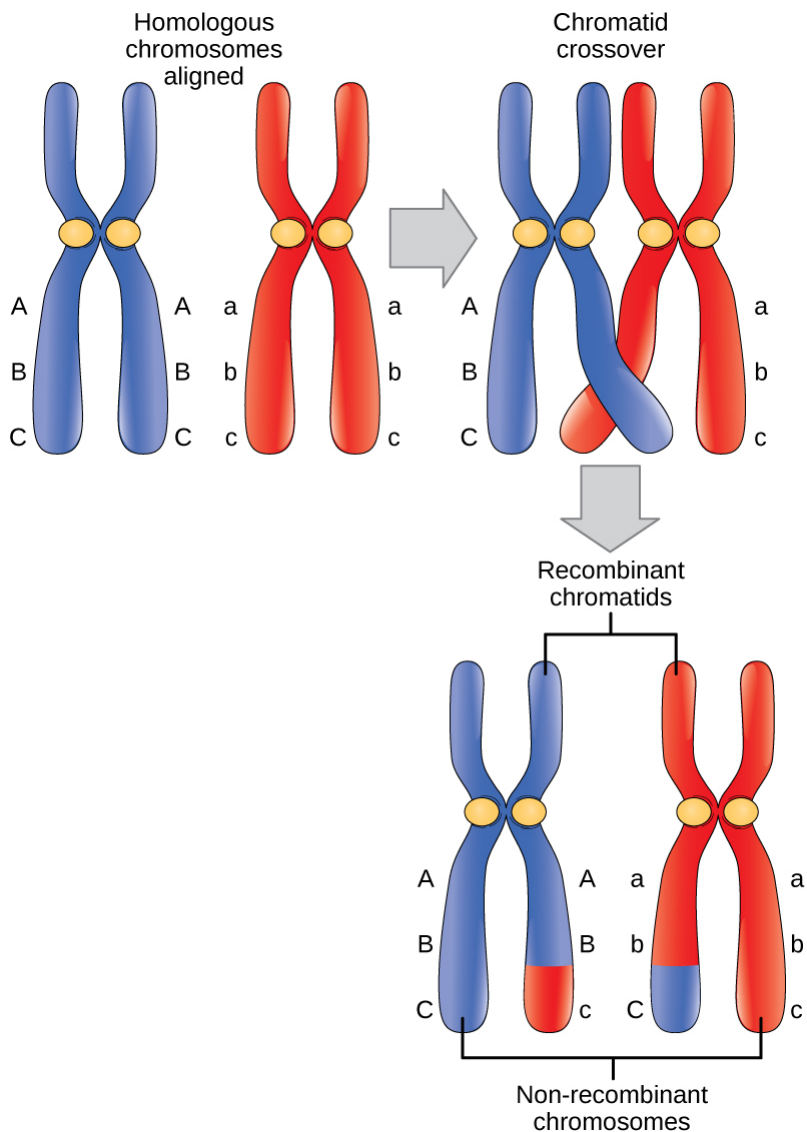


Located at intervals along the synaptonemal

complex are large protein assemblies called **recombination nodules**. These assemblies mark the points of later chiasmata and mediate the multistep process of **crossover**—or genetic recombination—between the nonsister chromatids. Near the recombination nodule, the double-stranded DNA of each chromatid is cleaved, the cut ends are modified, and a new connection is made between the nonsister chromatids. As prophase I progresses, the synaptonemal complex begins to break down and the chromosomes begin to condense. When the synaptonemal complex is gone, the homologous chromosomes remain attached to each other at the centromere and at chiasmata. The chiasmata remain until anaphase I. The number of chiasmata varies according to the species and the length of the chromosome. There must be at least one chiasma per chromosome for proper separation of homologous chromosomes during meiosis I, but there may be as many as 25. Following crossover, the synaptonemal complex breaks down and the cohesin connection between homologous pairs is removed. At the end of prophase I, the pairs are held together only at the chiasmata ([\[link\]](#)). These pairs are called **tetrads** because the four sister chromatids of each pair of homologous chromosomes are now visible.

The crossover events are the first source of genetic variation in the nuclei produced by meiosis. A single crossover event between homologous nonsister

chromatids leads to a reciprocal exchange of equivalent DNA between a maternal chromosome and a paternal chromosome. When a recombinant sister chromatid is moved into a gamete cell it will carry some DNA from one parent and some DNA from the other parent. The recombinant chromatid has a combination of maternal and paternal genes that did not exist before the crossover. Crossover events can occur almost anywhere along the length of the synapsed chromosomes. Different cells undergoing meiosis will therefore produce different recombinant chromatids, with varying combinations of maternal and parental genes. Multiple crossovers in an arm of the chromosome have the same effect, exchanging segments of DNA to produce genetically recombined chromosomes.



Prometaphase I

The key event in prometaphase I is the attachment of the spindle fiber microtubules to the kinetochore proteins at the centromeres. Kinetochore proteins

are multiprotein complexes that bind the centromeres of a chromosome to the microtubules of the mitotic spindle. Microtubules grow from microtubule-organizing centers (MTOCs). In animal cells, MTOCs are centrosomes located at opposite poles of the cell. The microtubules from each pole move toward the middle of the cell and attach to one of the kinetochores of the two fused homologous chromosomes. Each member of the homologous pair attaches to a microtubule extending from opposite poles of the cell so that in the next phase, the microtubules can pull the homologous pair apart. A spindle fiber that has attached to a kinetochore is called a *kinetochore microtubule*. At the end of prometaphase I, each tetrad is attached to microtubules from both poles, with one homologous chromosome facing each pole. The homologous chromosomes are still held together at the chiasmata. In addition, the nuclear membrane has broken down entirely.

Metaphase I

During metaphase I, the homologous chromosomes are arranged at the **metaphase plate**—roughly in the midline of the cell, with the kinetochores facing opposite poles. The homologous pairs orient themselves randomly at the equator. For example, if the two homologous members of chromosome 1 are labeled *a* and *b*, then the chromosomes could line up a-b or b-a. This is important in determining the

genes carried by a gamete, as each will only receive one of the two homologous chromosomes. (Recall that after crossing over takes place, homologous chromosomes are not identical. They contain slight differences in their genetic information, causing each gamete to have a unique genetic makeup.)

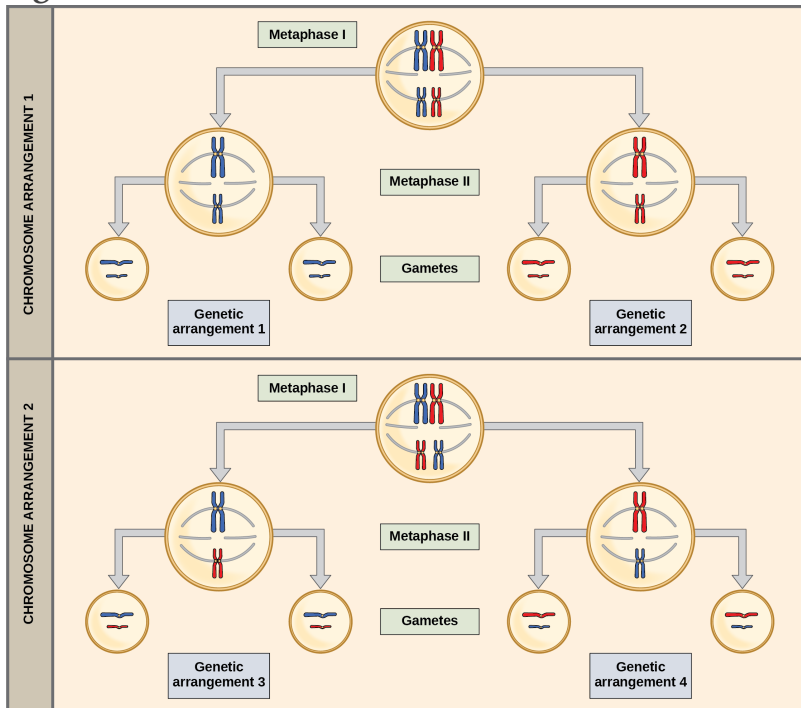
The randomness in the alignment of recombined chromosomes at the metaphase plate, coupled with the crossing over events between nonsister chromatids, are responsible for much of the genetic variation in the offspring. To clarify this further, remember that the homologous chromosomes of a sexually reproducing organism are originally inherited as two separate sets, one from each parent. Using humans as an example, one set of 23 chromosomes is present in the egg donated by the mother. The father provides the other set of 23 chromosomes in the sperm that fertilizes the egg. Every cell of the multicellular offspring has copies of the original two sets of homologous chromosomes. In prophase I of meiosis, the homologous chromosomes form the tetrads. In metaphase I, these pairs line up at the midway point between the two poles of the cell to form the metaphase plate. Because there is an equal chance that a microtubule fiber will encounter a maternally or paternally inherited chromosome, the arrangement of the tetrads at the metaphase plate is random. Thus, any maternally inherited chromosome may face either pole. Likewise, any paternally inherited

chromosome may also face either pole. *The orientation of each tetrad is independent of the orientation of the other 22 tetrads.*

This event—the *random* (or *independent*) assortment of homologous chromosomes at the metaphase plate—is the second mechanism that introduces variation into the gametes or spores. In each cell that undergoes meiosis, the arrangement of the tetrads is different. The number of variations is dependent on the number of chromosomes making up a set. There are two possibilities for orientation at the metaphase plate; the possible number of alignments therefore equals 2^n in a diploid cell, where n is the number of chromosomes per haploid set. Humans have 23 chromosome pairs, which results in over eight million (2^{23}) possible genetically-distinct gametes just from the random alignment of chromosomes at the metaphase plate. This number does not include the variability that was previously produced by crossing over between the nonsister chromatids. Given these two mechanisms, it is highly unlikely that any two haploid cells resulting from meiosis will have the same genetic composition ([\[link\]](#)).

To summarize, meiosis I creates genetically diverse gametes in two ways. First, during prophase I, crossover events between the nonsister chromatids of each homologous pair of chromosomes generate recombinant chromatids with new combinations of

maternal and paternal genes. Second, the random assortment of tetrads on the metaphase plate produces unique combinations of maternal and paternal chromosomes that will make their way into the gametes.



Anaphase I

In anaphase I, the microtubules pull the linked chromosomes apart. The sister chromatids remain tightly bound together at the centromere. The chiasmata are broken in anaphase I as the microtubules attached to the fused kinetochores pull the homologous chromosomes apart ([\[link\]](#)).

Telophase I and Cytokinesis

In telophase, the separated chromosomes arrive at opposite poles. The remainder of the typical telophase events may or may not occur, depending on the species. In some organisms, the chromosomes “decondense” and nuclear envelopes form around the separated sets of chromatids produced during telophase I. In other organisms, **cytokinesis**—the physical separation of the cytoplasmic components into two daughter cells—occurs without reformation of the nuclei. In nearly all species of animals and some fungi, cytokinesis separates the cell contents via a *cleavage furrow* (constriction of the actin ring that leads to cytoplasmic division). In plants, a *cell plate* is formed during cell cytokinesis by Golgi vesicles fusing at the metaphase plate. This cell plate will ultimately lead to the formation of cell walls that separate the two daughter cells.

Two haploid cells are the result of the first meiotic division of a diploid cell. The cells are haploid because at each pole, there is just one of each pair of the homologous chromosomes. Therefore, only one full set of the chromosomes is present. This is why the cells are considered haploid—there is only one chromosome set, *even though each chromosome still consists of two sister chromatids*. Recall that sister chromatids are merely duplicates of one of the two homologous chromosomes (except for changes that occurred during crossing over). In meiosis II, these

two sister chromatids will separate, creating four haploid daughter cells.

Link to Learning

Review the process of meiosis, observing how chromosomes align and migrate, at [Meiosis: An Interactive Animation](#).

The process of chromosome alignment differs between meiosis I and meiosis II. In prometaphase I, microtubules attach to the fused kinetochores of homologous chromosomes, and the homologous chromosomes are arranged at the midline of the cell (the metaphase plate) in metaphase I. In anaphase I, the homologous chromosomes separate. In prometaphase II, microtubules attach to the kinetochores of sister chromatids, and the sister chromatids are arranged at the midpoint of the cells in metaphase II. In anaphase II, the sister chromatids separate. An animal cell with a diploid number of four ($2n = 4$) proceeds through the stages of meiosis to form four haploid daughter cells.

Meiosis II

In some species, cells enter a brief interphase, or

interkinesis, before entering meiosis II. Interkinesis lacks an S phase, so chromosomes are not duplicated. The two cells produced in meiosis I go through the events of meiosis II in synchrony. During meiosis II, the sister chromatids within the two daughter cells separate, forming four new haploid gametes. The mechanics of meiosis II are similar to mitosis, except that each dividing cell has only one set of homologous chromosomes, each with two chromatids. Therefore, each cell has half the number of sister chromatids to separate out as a diploid cell undergoing mitosis. In terms of chromosomal content, cells at the start of meiosis II are similar to haploid cells in G₂, preparing to undergo mitosis.

Prophase II

If the chromosomes decondensed in telophase I, they condense again. If nuclear envelopes were formed, they fragment into vesicles. The MTOCs that were duplicated during interkinesis move away from each other toward opposite poles, and new spindles are formed.

Prometaphase II

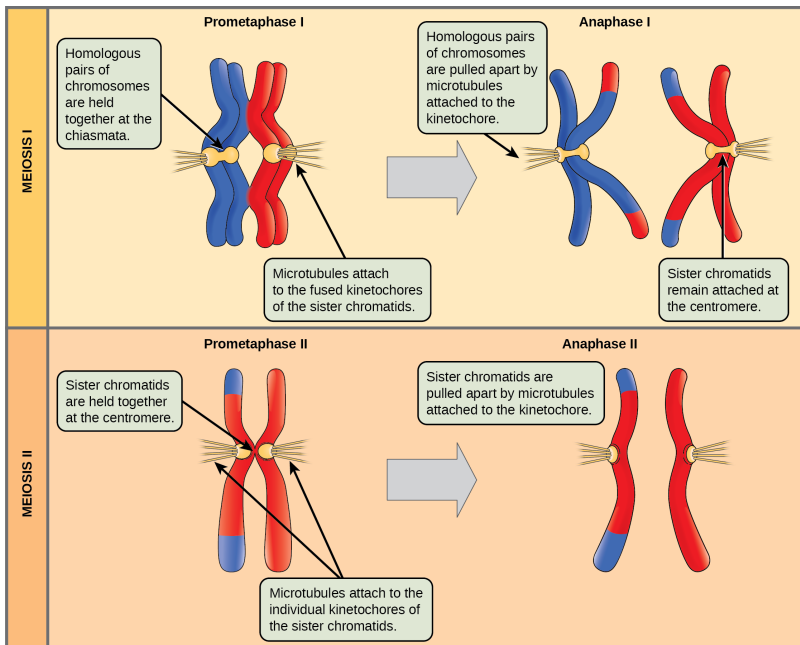
The nuclear envelopes are completely broken down, and the spindle is fully formed. Each sister chromatid forms an individual kinetochore that attaches to microtubules from opposite poles.

Metaphase II

The sister chromatids are maximally condensed and aligned at the equator of the cell.

Anaphase II

The sister chromatids are pulled apart by the kinetochore microtubules and move toward opposite poles. Nonkinetochore microtubules elongate the cell.



Telophase II and Cytokinesis

The chromosomes arrive at opposite poles and begin to decondense. Nuclear envelopes form around the

chromosomes. If the parent cell was diploid, as is most commonly the case, then cytokinesis now separates the two cells into four unique haploid cells. The cells produced are genetically unique because of the random assortment of paternal and maternal homologs and because of the recombination of maternal and paternal segments of chromosomes (with their sets of genes) that occurs during crossover. The entire process of meiosis is outlined in [\[link\]](#).

	Stage	Event	Outcome
INTERPHASE	S phase		Chromosomes are duplicated during interphase. The resulting sister chromatids are held together at the centromere. The centrosomes are also duplicated.
	Prophase I		Chromosomes condense, and the nuclear envelope fragments. Homologous chromosomes bind firmly together along their length, forming a tetrad. Chiasmata form between non sister chromatids. Crossing over occurs at the chiasmata. Spindle fibers emerge from the centrosomes.
MEIOSIS I	Prometaphase I		Homologous chromosomes are attached to spindle microtubules at the fused kinetochore shared by the sister chromatids. Chromosomes continue to condense, and the nuclear envelope completely disappears.
	Metaphase I		Homologous chromosomes randomly assemble at the metaphase plate, where they have been maneuvered into place by the microtubules.
	Anaphase I		Spindle microtubules pull the homologous chromosomes apart. The sister chromatids are still attached at the centromere.
	Telophase I and Cytokinesis		Sister chromatids arrive at the poles of the cell and begin to decondense. A nuclear envelope forms around each nucleus, and the cytoplasm is divided by a cleavage furrow. The result is two haploid cells. Each cell contains one duplicated copy of each homologous chromosome pair.
MEIOSIS II	Prophase II		Sister chromatids condense. A new spindle begins to form. The nuclear envelope starts to fragment.
	Prometaphase II		The nuclear envelope disappears, and the spindle fibers engage the individual kinetochores on the sister chromatids.
	Metaphase II		Sister chromatids line up at the metaphase plate.
	Anaphase II		Sister chromatids are pulled apart by the shortening of the kinetochore microtubules. Non kinetochore microtubules lengthen the cell.
	Telophase II and Cytokinesis		Chromosomes arrive at the poles of the cell and decondense. Nuclear envelopes surround the four nuclei. Cleavage furrows divide the two cells into four haploid cells.

Meiosis and mitosis are both preceded by one cycle of DNA replication; however, meiosis includes two nuclear divisions. The four daughter cells resulting from meiosis are haploid and genetically distinct. The daughter cells resulting from mitosis are diploid and identical to the parent cell.

Comparing Meiosis and Mitosis

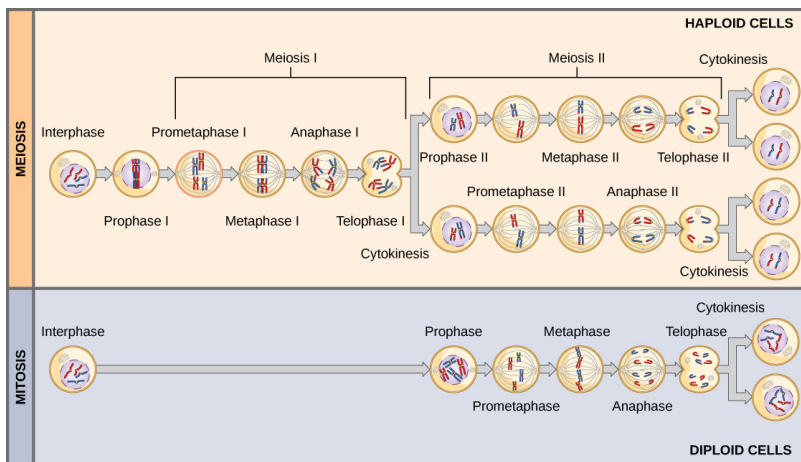
Mitosis and meiosis are both forms of division of the nucleus in eukaryotic cells. They share some similarities, but also exhibit a number of important and distinct differences that lead to very different outcomes ([\[link\]](#)). Mitosis is a single nuclear division that results in two nuclei that are usually partitioned into two new cells. The nuclei resulting from a mitotic division are genetically identical to the original nucleus. They have the same number of sets of chromosomes: one set in the case of haploid cells and two sets in the case of diploid cells. In contrast, meiosis consists of two nuclear divisions resulting in four nuclei that are usually partitioned into four new, genetically distinct cells. The four nuclei produced during meiosis are not genetically identical, and they contain one chromosome set only. This is half the number of chromosome sets in the original cell, which is diploid.

The main differences between mitosis and meiosis occur in meiosis I, which is a very different nuclear division than mitosis. In meiosis I, the homologous chromosome pairs physically meet and are bound together with the synaptonemal complex. Following this, the chromosomes develop chiasmata and undergo crossover between nonsister chromatids. In the end, the chromosomes line up along the metaphase plate as tetrads—with kinetochore fibers from opposite spindle poles attached to each

kinetochore of a homolog to form a tetrad. *All of these events occur only in meiosis I.*

When the chiasmata resolve and the tetrad is broken up with the homologs moving to one pole or another, the ploidy level—the number of sets of chromosomes in each future nucleus—has been reduced from two to one. For this reason, meiosis I is referred to as a **reductional division**. There is no such reduction in ploidy level during mitosis.

Meiosis II is analogous to a mitotic division. In this case, the duplicated chromosomes (only one set of them) line up on the metaphase plate with divided kinetochores attached to kinetochore fibers from opposite poles. During anaphase II, as in mitotic anaphase, the kinetochores divide and one sister chromatid—now referred to as a chromosome—is pulled to one pole while the other sister chromatid is pulled to the other pole. If it were not for the fact that there had been crossover, the two products of each individual meiosis II division would be identical (as in mitosis). Instead, they are different because there has always been at least one crossover per chromosome. Meiosis II is not a reduction division because although there are fewer copies of the genome in the resulting cells, there is still one set of chromosomes, as there was at the end of meiosis I.



						OUTCOME
PROCESS	DNA synthesis	Synapsis of homologous chromosomes	Crossover	Homologous chromosomes line up at metaphase plate	Sister chromatids line up at metaphase plate	Number and genetic composition of daughter cells
MEIOSIS	Occurs in S phase of interphase	During prophase I	During prophase I	During metaphase I	During metaphase II	Four haploid cells at the end of meiosis II
MITOSIS	Occurs in S phase of interphase	Does not occur in mitosis	Does not occur in mitosis	Does not occur in mitosis	During metaphase	Two diploid cells at the end of mitosis

Evolution Connection

The Mystery of the Evolution of Meiosis

Some characteristics of organisms are so widespread and fundamental that it is sometimes difficult to remember that they evolved like other simple traits. Meiosis is such an extraordinarily complex series of cellular events that biologists have had trouble testing hypotheses concerning how it may have evolved. Although meiosis is inextricably entwined with sexual reproduction and its advantages and disadvantages, it is important to separate the questions of the evolution of meiosis

and the evolution of sex, because early meiosis may have been advantageous for different reasons than it is now. Thinking outside the box and imagining what the early benefits from meiosis might have been is one approach to uncovering how it may have evolved.

Meiosis and mitosis share obvious cellular processes, and it makes sense that meiosis evolved from mitosis. The difficulty lies in the clear differences between meiosis I and mitosis. Adam Wilkins and Robin Holliday^[footnote] summarized the unique events that needed to occur for the evolution of meiosis from mitosis. These steps are homologous chromosome pairing and synapsis, crossover exchanges, sister chromatids remaining attached during anaphase, and suppression of DNA replication in interphase. They argue that the first step is the hardest and most important and that understanding how it evolved would make the evolutionary process clearer. They suggest genetic experiments that might shed light on the evolution of synapsis.

Adam S. Wilkins and Robin Holliday, “The Evolution of Meiosis from Mitosis,” *Genetics* 181 (2009): 3–12.

There are other approaches to understanding the evolution of meiosis in progress. Different forms of meiosis exist in single-celled protists. Some appear to be simpler or more “primitive” forms of meiosis. Comparing the meiotic divisions of different protists may shed light on the evolution of meiosis.

Marilee Ramesh and colleagues^[footnote] compared the genes involved in meiosis in protists to understand when and where meiosis might have evolved. Although research is still ongoing, recent scholarship into meiosis in protists suggests that some aspects of meiosis may have evolved later than others. This kind of genetic comparison can tell us what aspects of meiosis are the oldest and what cellular processes they may have borrowed from in earlier cells.

Marilee A. Ramesh, Shehre-Banoo Malik and John M. Logsdon, Jr, “A Phylogenetic Inventory of Meiotic Genes: Evidence for Sex in *Giardia* and an Early Eukaryotic Origin of Meiosis,” *Current Biology* 15 (2005):185–91.

Link to Learning

Click through the steps of this interactive animation to compare the meiotic process of cell division to that of mitosis at [How Cells Divide](#).

Section Summary

Sexual reproduction requires that organisms

produce cells that can fuse during fertilization to produce offspring. In most animals, meiosis is used to produce haploid eggs and sperm from diploid parent cells so that the fusion of an egg and sperm produces a diploid zygote. As with mitosis, DNA replication occurs prior to meiosis during the S-phase of the cell cycle so that each chromosome becomes a pair of sister chromatids. In meiosis, there are two rounds of nuclear division resulting in four nuclei and usually four daughter cells, each with half the number of chromosomes as the parent cell. The first division separates homologs, and the second—like mitosis—separates chromatids into individual chromosomes. Meiosis generates variation in the daughter nuclei during crossover in prophase I as well as during the random alignment of tetrads at metaphase I. The cells that are produced by meiosis are genetically unique.

Meiosis and mitosis share similar processes, but have distinct outcomes. Mitotic divisions are single nuclear divisions that produce genetically identical daughter nuclei (i.e., each daughter nucleus has the same number of chromosome sets as the original cell). In contrast, meiotic divisions include two nuclear divisions that ultimately produce four genetically different daughter nuclei that have only one chromosome set (instead of the two sets of chromosomes in the parent cell). The main differences between the two nuclear division processes take place during the first division of

meiosis: homologous chromosomes pair, crossover, and exchange homologous nonsister chromatid segments. The homologous chromosomes separate into different nuclei during meiosis I, causing a *reduction of ploidy level in the first division*. The second division of meiosis is similar to a mitotic division, except that the daughter cells do not contain identical genomes because of crossover and chromosome recombination in prophase I.

Review Questions

Meiosis usually produces _____ daughter cells.

1. two haploid
2. two diploid
3. four haploid
4. four diploid

C

What structure is most important in forming the tetrads?

1. centromere
2. synaptonemal complex
3. chiasma

4. kinetochore

B

At which stage of meiosis are sister chromatids separated from each other?

1. prophase I
 2. prophase II
 3. anaphase I
 4. anaphase II
-

D

At metaphase I, homologous chromosomes are connected only at what structures?

1. chiasmata
 2. recombination nodules
 3. microtubules
 4. kinetochores
-

A

Which of the following is *not* true in regard to crossover?

1. Spindle microtubules guide the transfer of DNA across the synaptonemal complex.
 2. Nonsister chromatids exchange genetic material.
 3. Chiasmata are formed.
 4. Recombination nodules mark the crossover point.
-

C

What phase of mitotic interphase is missing from meiotic interkinesis?

1. G₀ phase
 2. G₁ phase
 3. S phase
 4. G₂ phase
-

C

The part of meiosis that is similar to mitosis is _____.

1. meiosis I
 2. anaphase I
 3. meiosis II
 4. interkinesis
-

C

If a muscle cell of a typical organism has 32 chromosomes, how many chromosomes will be in a gamete of that same organism?

1. 8
2. 16
3. 32
4. 64

B

Which statement best describes the genetic content of the two daughter cells in prophase II of meiosis?

1. haploid with one copy of each gene
2. haploid with two copies of each gene
3. diploid with two copies of each gene
4. diploid with four copies of each gene

B

The pea plants used in Mendel's genetic inheritance studies were diploid, with 14 chromosomes in somatic cells. Assuming no

crossing over events occur, how many unique gametes could one pea plant produce?

1. 28
2. 128
3. 196
4. 16,384

B

How do telophase I and telophase II differ during meiosis in animal cells?

1. Cells remain diploid at the end of telophase I, but are haploid at the end of telophase II.
2. Daughter cells form a cell plate to divide during telophase I, but divide by cytokinesis during telophase II.
3. Cells enter interphase after telophase I, but not after telophase II.
4. Chromosomes can remain condensed at the end of telophase I, but decondense after telophase II.

D

Critical Thinking Questions

Describe the process that results in the formation of a tetrad.

During the meiotic interphase, each chromosome is duplicated. The sister chromatids that are formed during synthesis are held together at the centromere region by cohesin proteins. All chromosomes are attached to the nuclear envelope by their tips. As the cell enters prophase I, the nuclear envelope begins to fragment and the proteins holding homologous chromosomes locate each other. The four sister chromatids align lengthwise, and a protein lattice called the synaptonemal complex is formed between them to bind them together. The synaptonemal complex facilitates crossover between nonsister chromatids, which is observed as chiasmata along the length of the chromosome. As prophase I progresses, the synaptonemal complex breaks down and the sister chromatids become free, except where they are attached by chiasmata. At this stage, the four chromatids are visible in each homologous pairing and are called a tetrad.

Explain how the random alignment of

homologous chromosomes during metaphase I contributes to the variation in gametes produced by meiosis.

Random alignment leads to new combinations of traits. The chromosomes that were originally inherited by the gamete-producing individual came equally from the egg and the sperm. In metaphase I, the duplicated copies of these maternal and paternal homologous chromosomes line up across the center of the cell. The orientation of each tetrad is random. There is an equal chance that the maternally derived chromosomes will be facing either pole. The same is true of the paternally derived chromosomes. The alignment should occur differently in almost every meiosis. As the homologous chromosomes are pulled apart in anaphase I, any combination of maternal and paternal chromosomes will move toward each pole. The gametes formed from these two groups of chromosomes will have a mixture of traits from the individual's parents. Each gamete is unique.

What is the function of the fused kinetochore found on sister chromatids in prometaphase I?

In metaphase I, the homologous chromosomes

line up at the metaphase plate. In anaphase I, the homologous chromosomes are pulled apart and move to opposite poles. Sister chromatids are not separated until meiosis II. The fused kinetochore formed during meiosis I ensures that each spindle microtubule that binds to the tetrad will attach to both sister chromatids.

In a comparison of the stages of meiosis to the stages of mitosis, which stages are unique to meiosis and which stages have the same events in both meiosis and mitosis?

All of the stages of meiosis I, except possibly telophase I, are unique because homologous chromosomes are separated, not sister chromatids. In some species, the chromosomes do not decondense and the nuclear envelopes do not form in telophase I. All of the stages of meiosis II have the same events as the stages of mitosis, with the possible exception of prophase II. In some species, the chromosomes are still condensed and there is no nuclear envelope. Other than this, all processes are the same.

Why would an individual with a mutation that prevented the formation of recombination nodules be considered less fit than other members of its species?

The chromosomes of the individual cannot cross over during meiosis if the individual cannot make recombination nodules. This limits the genetic diversity of the individual's gametes to what occurs during independent assortment, with all daughter cells receiving complete maternal or paternal chromatids. An individual who cannot produce diverse offspring is considered less fit than individuals who do produce diverse offspring.

Does crossing over occur during prophase II? From an evolutionary perspective, why is this advantageous?

Crossing over does not occur during prophase II; it only occurs during prophase I. In prophase II, there are still two copies of each gene, but they are on sister chromatids within a single chromosome (rather than homologous chromosomes as in prophase I). Therefore, any crossover event would still produce two identical chromatids. Because it is advantageous to avoid wasting energy on events that will not increase genetic diversity, crossing over does not occur.

Glossary

chiasmata

(singular, *chiasma*) the structure that forms at the crossover points after genetic material is exchanged

cohesin

proteins that form a complex that seals sister chromatids together at their centromeres until anaphase II of meiosis

crossover

exchange of genetic material between nonsister chromatids resulting in chromosomes that incorporate genes from both parents of the organism

fertilization

union of two haploid cells from two individual organisms

interkinesis

(also, *interphase II*) brief period of rest between meiosis I and meiosis II

meiosis

a nuclear division process that results in four haploid cells

meiosis I

first round of meiotic cell division; referred to as *reduction division* because the ploidy level is reduced from diploid to haploid

meiosis II

second round of meiotic cell division following meiosis I; sister chromatids are separated into individual chromosomes, and the result is four unique haploid cells

recombination nodules

protein assemblies formed on the synaptonemal complex that mark the points of crossover events and mediate the multistep process of genetic recombination between nonsister chromatids

reduction division

nuclear division that produces daughter nuclei each having one-half as many chromosome sets as the parental nucleus; meiosis I is a reduction division

somatic cell

all the cells of a multicellular organism except the gametes or reproductive cells

spore

haploid cell that can produce a haploid multicellular organism or can fuse with another spore to form a diploid cell

synapsis

formation of a close association between homologous chromosomes during prophase I

synaptonemal complex

protein lattice that forms between
homologous chromosomes during prophase I,
supporting crossover

tetrad

two duplicated homologous chromosomes
(four chromatids) bound together by
chiasmata during prophase I

Sexual Reproduction

By the end of this section, you will be able to do the following:

- Explain that meiosis and sexual reproduction are highly evolved traits
- Identify variation among offspring as a potential evolutionary advantage of sexual reproduction
- Describe the three different life-cycle types among sexually reproducing multicellular organisms.

Sexual reproduction was likely an early evolutionary innovation after the appearance of eukaryotic cells. It appears to have been very successful because most eukaryotes are able to reproduce sexually and, in many animals, it is the only mode of reproduction. And yet, scientists also recognize some real disadvantages to sexual reproduction. On the surface, creating offspring that are genetic clones of the parent appears to be a better system. If the parent organism is successfully occupying a habitat, offspring with the same traits should be similarly successful. There is also the obvious benefit to an organism that can produce offspring whenever circumstances are favorable by asexual budding, fragmentation, or by producing eggs asexually. These methods of reproduction do not require another organism of the opposite sex. Indeed, some organisms that lead a solitary lifestyle

have retained the ability to reproduce asexually. In addition, in asexual populations, every individual is capable of reproduction. In sexual populations, the males are not producing the offspring themselves, so hypothetically an asexual population could grow twice as fast.

However, multicellular organisms that exclusively depend on asexual reproduction are exceedingly rare. Why are meiosis and sexual reproductive strategies so common? These are important (and as yet unanswered) questions in biology, even though they have been the focus of much research beginning in the latter half of the 20th century. There are several possible explanations, one of which is that the variation that sexual reproduction creates among offspring is very important to the survival and reproduction of the population. Thus, on average, a sexually reproducing population will leave more descendants than an otherwise similar asexually reproducing population. The only source of variation in asexual organisms is mutation. Mutations that take place during the formation of germ cell lines are also the ultimate source of variation in sexually reproducing organisms. However, in contrast to mutation during asexual reproduction, the mutations during sexual reproduction can be continually reshuffled from one generation to the next when different parents combine their unique genomes and the genes are mixed into different combinations by crossovers

during prophase I and random assortment at metaphase I.

Evolution Connection

The Red Queen Hypothesis

Genetic variation is the outcome of sexual reproduction, but why are ongoing variations necessary, even under seemingly stable environmental conditions? Enter the Red Queen hypothesis, first proposed by Leigh Van Valen in 1973.^[footnote] The concept was named in reference to the Red Queen's race in Lewis Carroll's book, *Through the Looking-Glass*.

Leigh Van Valen, "A New Evolutionary Law," *Evolutionary Theory* 1 (1973): 1–30

All species **coevolve** (evolve together) with other organisms. For example, predators evolve with their prey, and parasites evolve with their hosts. Each tiny advantage gained by favorable variation gives a species a reproductive edge over close competitors, predators, parasites, or even prey. However, survival of any given genotype or phenotype in a population is dependent on the reproductive fitness of other genotypes or phenotypes within a given species. The only method that will allow a coevolving species to maintain its own share of the resources is to also *continually improve its fitness* (the capacity of the members to produce more reproductively viable

offspring relative to others within a species). As one species gains an advantage, this increases selection on the other species; they must also develop an advantage or they will be outcompeted. No single species progresses too far ahead because genetic variation among the progeny of sexual reproduction provides all species with a mechanism to improve rapidly. Species that cannot keep up become extinct. The Red Queen's catchphrase was, "It takes all the running you can do to stay in the same place." This is an apt description of coevolution between competing species.

In animals, sexually reproducing adults form haploid gametes from diploid germ cells. Fusion of the gametes gives rise to a fertilized egg cell, or zygote. The zygote will undergo multiple rounds of mitosis to produce a multicellular offspring. The germ cells are generated early in the development of the zygote. Plants have a life cycle that alternates between a multicellular haploid organism and a multicellular diploid organism. In some plants, such as ferns, both the haploid and diploid plant stages are free-living. The diploid plant is called a sporophyte because it produces haploid spores by meiosis. The spores develop into multicellular, haploid plants that are called *gametophytes* because they produce gametes. The gametes of two

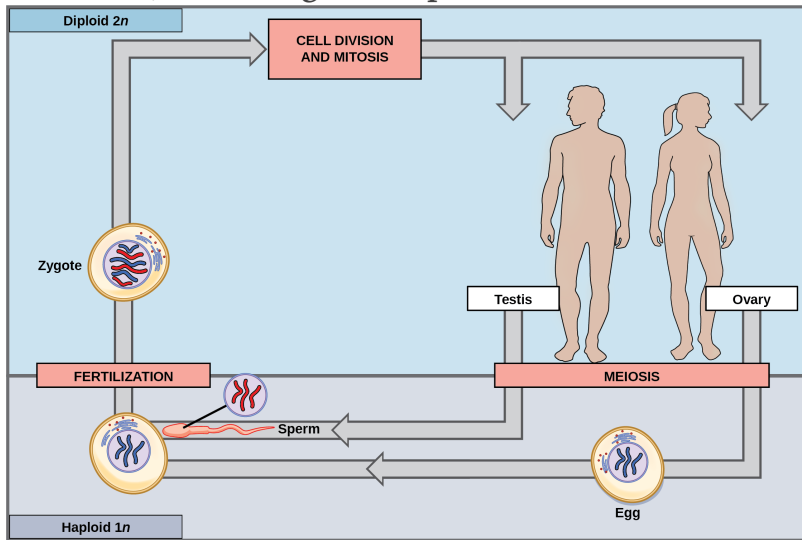
individuals will fuse to form a diploid zygote that becomes the sporophyte. (credit “fern”: modification of work by Cory Zanker; credit “sporangia”: modification of work by "Obsidian Soul"/Wikimedia Commons; credit “gametophyte and sporophyte”: modification of work by “Vlmastra”/Wikimedia Commons)

Life Cycles of Sexually Reproducing Organisms

Fertilization and meiosis alternate in sexual **life cycles**. What happens between these two events depends on the organism’s “reproductive strategy.” The process of meiosis reduces the chromosome number by half. Fertilization, the joining of two haploid gametes, restores the diploid condition. Some organisms have a multicellular diploid stage that is most obvious and only produce haploid reproductive cells. Animals, including humans, have this type of life cycle. Other organisms, such as fungi, have a multicellular haploid stage that is most obvious. Plants and some algae have alternation of generations, in which they have multicellular diploid and haploid life stages that are apparent to different degrees depending on the group.

Nearly all animals employ a diploid-dominant life-cycle strategy in which the only haploid cells produced by the organism are the gametes. Early in the development of the embryo, specialized diploid

cells, called **germ cells**, are produced within the gonads (such as the testes and ovaries). Germ cells are capable of mitosis to perpetuate the germ cell line and meiosis to produce haploid gametes. Once the haploid gametes are formed, they lose the ability to divide again. There is no multicellular haploid life stage. Fertilization occurs with the fusion of two gametes, usually from different individuals, restoring the diploid state ([\[link\]](#)).

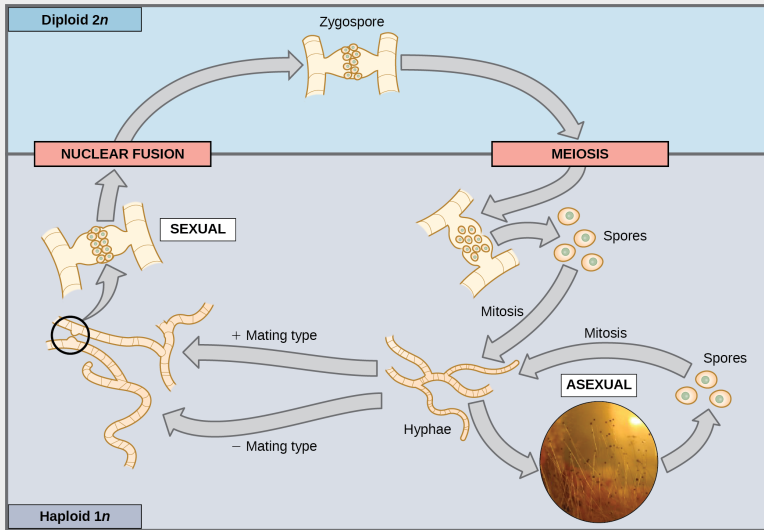


Most fungi and algae employ a life-cycle type in which the “body” of the organism—the ecologically important part of the life cycle—is haploid. The haploid cells that make up the tissues of the dominant multicellular stage are formed by mitosis. During sexual reproduction, specialized haploid cells from two individuals—designated the (+) and (–) mating types—join to form a diploid zygote. The zygote immediately undergoes meiosis to form

four haploid cells called *spores*. Although these spores are haploid like the “parents,” they contain a new genetic combination from two parents. The spores can remain dormant for various time periods. Eventually, when conditions are favorable, the spores form multicellular haploid structures through many rounds of mitosis ([\[link\]](#)).

Visual Connection

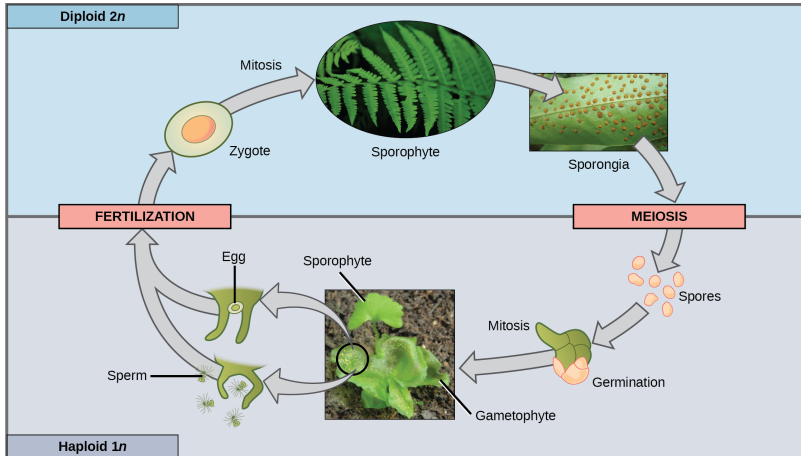
Fungi, such as black bread mold (*Rhizopus nigricans*), have a haploid multicellular stage that produces specialized haploid cells by mitosis that fuse to form a diploid zygote. The haploid multicellular stage produces specialized haploid cells by mitosis that fuse to form a diploid zygote. The zygote undergoes meiosis to produce haploid spores. Each spore gives rise to a multicellular haploid organism by mitosis. Above, different mating hyphae types (denoted as + and –) join to form a zygosporangium through nuclear fusion. (credit “zygomycota” micrograph: modification of work by “Fanaberka”/Wikimedia Commons)



If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

The third life-cycle type, employed by some algae and all plants, is a blend of the haploid-dominant and diploid-dominant extremes. Species with alternation of generations have both haploid and diploid multicellular organisms as part of their life cycle. The haploid multicellular plants are called **gametophytes**, because they produce gametes from specialized cells. Meiosis is not directly involved in the production of gametes in this case, because the organism that produces the gametes is already haploid. Fertilization between the gametes forms a diploid zygote. The zygote will undergo many rounds of mitosis and give rise to a diploid

multicellular plant called a **sporophyte**. Specialized cells of the sporophyte will undergo meiosis and produce haploid spores. The spores will subsequently develop into the gametophytes ([link]).



Although all plants utilize some version of the alternation of generations, the relative size of the sporophyte and the gametophyte and the relationship between them vary greatly. In plants such as moss, the gametophyte organism is the free-living plant and the sporophyte is physically dependent on the gametophyte. In other plants, such as ferns, both the gametophyte and sporophyte plants are free-living; however, the sporophyte is much larger. In seed plants, such as magnolia trees and daisies, the gametophyte is composed of only a few cells and, in the case of the female gametophyte, is completely retained within the

sporophyte.

Sexual reproduction takes many forms in multicellular organisms. The fact that nearly every multicellular organism on Earth employs sexual reproduction is strong evidence for the benefits of producing offspring with unique gene combinations, though there are other possible benefits as well.

Section Summary

Nearly all eukaryotes undergo sexual reproduction. The variation introduced into the reproductive cells by meiosis provides an important advantage that has made sexual reproduction evolutionarily successful. Meiosis and fertilization alternate in sexual life cycles. The process of meiosis produces unique reproductive cells called gametes, which have half the number of chromosomes as the parent cell. When two haploid gametes fuse, this restores the diploid condition in the new zygote. Thus, most sexually reproducing organisms alternate between haploid and diploid stages. However, the ways in which reproductive cells are produced and the timing between meiosis and fertilization vary greatly.

Visual Connection Questions

[\[link\]](#) If a mutation occurs so that a fungus is no longer able to produce a minus mating type, will it still be able to reproduce?

[\[link\]](#) Yes, it will be able to reproduce asexually.

Review Questions

What is a likely evolutionary advantage of sexual reproduction over asexual reproduction?

1. Sexual reproduction involves fewer steps.
 2. There is a lower chance of using up the resources in a given environment.
 3. Sexual reproduction results in variation in the offspring.
 4. Sexual reproduction is more cost-effective.
-

C

Which type of life cycle has both a haploid and diploid multicellular stage?

1. asexual life cycles
 2. most animal life cycles
 3. most fungal life cycles
 4. alternation of generations
-

D

What is the ploidy of the most conspicuous form of most fungi?

1. diploid
 2. haploid
 3. alternation of generations
 4. asexual
-

B

A diploid, multicellular life-cycle stage that gives rise to haploid cells by meiosis is called a _____.

1. sporophyte
 2. gametophyte
 3. spore
 4. gamete
-

A

Hydras and jellyfish both live in a freshwater lake that is slowly being acidified by the runoff from a chemical plant built upstream. Which population is predicted to be better able to cope with the changing environment?

1. jellyfish
2. hydra
3. The populations will be equally able to cope.
4. Both populations will die.

A

Many farmers are worried about the decreasing genetic diversity of plants associated with generations of artificial selection and inbreeding. Why is limiting random sexual reproduction of food crops concerning?

1. Mutations during asexual reproduction decrease plant fitness.
 2. Consumers do not trust identical-appearing produce.
 3. Larger portions of the plant populations are susceptible to the same diseases.
 4. Spores are not viable in an agricultural setting.
-

Critical Thinking Questions

List and briefly describe the three processes that lead to variation in offspring with the same parents.

a. Crossover occurs in prophase I between nonsister homologous chromosomes. Segments of DNA are exchanged between maternally derived and paternally derived chromosomes, and new gene combinations are formed. b. Random alignment during metaphase I leads to gametes that have a mixture of maternal and paternal chromosomes. c. Fertilization is random, in that any two gametes can fuse.

Animals and plants both have diploid and haploid cells. How does the animal life cycle differ from the alternation of generations exhibited by plants?

Nearly all animals employ a diploid-dominant life-cycle strategy; only the gametes are

haploid. Once the haploid gametes are formed, they lose the ability to divide again. There is no multicellular haploid life stage. Plants, in contrast, have a blend of the haploid-dominant and diploid-dominant cycles -- they have both haploid and diploid multicellular organisms as part of their life cycle. The diploid plant is called a sporophyte because it produces haploid spores by meiosis. The spores develop into multicellular, haploid plants that are called gametophytes because they produce gametes.

Explain why sexual reproduction is beneficial to a population but can be detrimental to an individual offspring.

Sexual reproduction increases the genetic variation within the population, because new individuals are made by randomly combining genetic material from two parents. Because only fit individuals reach sexual maturity and reproduce, the overall population tends toward increasing fitness in its environment. However, there is always a possibility that the random combination creating the offspring's genome will actually produce an organism less fit for the environment than its parents were.

How does the role of meiosis in gamete

production differ between organisms with a diploid-dominant life cycle and organisms with an alternation of generations life cycle?

Organisms with a diploid-dominant life cycle make haploid gametes by meiosis, while all their somatic cells are diploid. Organisms with an alternation of generations life cycle make gametes during their haploid life stage, so the chromosome number does not need to be reduced, and meiosis is not involved.

How do organisms with haploid-dominant life cycles ensure continued genetic diversification in offspring without using a meiotic process to make gametes?

Haploid-dominant organisms undergo sexual reproduction by making a diploid zygote. The cells that make the gametes are derived from haploid cells, but the + and – mating types that produce the zygote are randomly combined. The zygote also undergoes meiosis to return to the haploid stage, so multiple steps add genetic diversity to haploid-dominant organisms.

Glossary

alternation of generations

life-cycle type in which the diploid and haploid stages alternate

gametophyte

a multicellular haploid life-cycle stage that produces gametes

germ cells

specialized cell line that produces gametes, such as eggs or sperm

life cycle

the sequence of events in the development of an organism and the production of cells that produce offspring

sporophyte

a multicellular diploid life-cycle stage that produces haploid spores by meiosis

Introduction

class = "introduction" Experimenting with thousands of garden peas, Mendel uncovered the fundamentals of genetics. (credit: modification of work by Jerry Kirkhart)



Genetics is the study of heredity. Johann Gregor Mendel set the framework for genetics long before chromosomes or genes had been identified, at a time when meiosis was not well understood. Mendel selected a simple biological system and conducted methodical, quantitative analyses using large sample sizes. Because of Mendel's work, the fundamental principles of heredity were revealed. We now know that genes, carried on chromosomes, are the basic functional units of heredity with the capability to be replicated, expressed, or mutated. Today, the postulates put forth by Mendel form the basis of classical, or Mendelian, genetics. Not all genes are transmitted from parents to offspring according to Mendelian genetics, but Mendel's experiments serve

as an excellent starting point for thinking about inheritance.

Mendel's Experiments and the Laws of Probability

By the end of this section, you will be able to do the following:

- Describe the scientific reasons for the success of Mendel's experimental work
- Describe the expected outcomes of monohybrid crosses involving dominant and recessive alleles
- Apply the sum and product rules to calculate probabilities

Johann Gregor Mendel is considered the father of genetics.



Johann Gregor Mendel (1822–1884) ([\[link\]](#)) was a lifelong learner, teacher, scientist, and man of faith. As a young adult, he joined the Augustinian Abbey of St. Thomas in Brno in what is now the Czech Republic. Supported by the monastery, he taught physics, botany, and natural science courses at the secondary and university levels. In 1856, he began a decade-long research pursuit involving inheritance patterns in honeybees and plants, ultimately settling on pea plants as his primary **model system** (a system with convenient characteristics used to study a specific biological phenomenon to be applied to other systems). In 1865, Mendel presented the

results of his experiments with nearly 30,000 pea plants to the local Natural History Society. He demonstrated that traits are transmitted from parents to offspring independently of other traits and in dominant and recessive patterns. In 1866, he published his work, *Experiments in Plant Hybridization*,^[footnote] in the proceedings of the Natural History Society of Brünn.

Johann Gregor Mendel, *Versuche über Pflanzenhybriden Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr, 1865 Abhandlungen*, 3–47. [for English translation see <http://www.mendelweb.org/Mendel.plain.html>]

Mendel's work went virtually unnoticed by the scientific community, which believed, incorrectly, that the process of inheritance involved a blending of parental traits that produced an intermediate physical appearance in offspring. The **blending theory of inheritance** asserted that the original parental traits were lost or absorbed by the blending in the offspring, but we now know that this is not the case. This hypothetical process appeared to be correct because of what we know now as continuous variation. **Continuous variation** results from the action of many genes to determine a characteristic like human height. Offspring appear to be a “blend” of their parents' traits.

Instead of continuous characteristics, Mendel worked with traits that were inherited in distinct

classes (specifically, violet versus white flowers); this is referred to as **discontinuous variation**. Mendel's choice of these kinds of traits allowed him to see experimentally that the traits were not blended in the offspring, nor were they absorbed, but rather that they kept their distinctness and could be passed on. In 1868, Mendel became abbot of the monastery and exchanged his scientific pursuits for his pastoral duties. He was not recognized for his extraordinary scientific contributions during his lifetime. In fact, it was not until 1900 that his work was rediscovered, reproduced, and revitalized by scientists on the brink of discovering the chromosomal basis of heredity.

Mendel's Model System

Mendel's seminal work was accomplished using the garden pea, *Pisum sativum*, to study inheritance. This species naturally self-fertilizes, such that pollen encounters ova within individual flowers. The flower petals remain sealed tightly until after pollination, preventing pollination from other plants. The result is highly inbred, or "true-breeding," pea plants. These are plants that always produce offspring that look like the parent. By experimenting with true-breeding pea plants, Mendel avoided the appearance of unexpected traits in offspring that might occur if the plants were not

true breeding. The garden pea also grows to maturity within one season, meaning that several generations could be evaluated over a relatively short time. Finally, large quantities of garden peas could be cultivated simultaneously, allowing Mendel to conclude that his results did not come about simply by chance.

In one of his experiments on inheritance patterns, Mendel crossed plants that were true-breeding for violet flower color with plants true-breeding for white flower color (the P generation). The resulting hybrids in the F₁ generation all had violet flowers. In the F₂ generation, approximately three quarters of the plants had violet flowers, and one quarter had white flowers.

Mendelian Crosses

Mendel performed **hybridizations**, which involve mating two true-breeding individuals that have different traits. In the pea, which is naturally self-pollinating, this is done by manually transferring pollen from the anther of a mature pea plant of one variety to the stigma of a separate mature pea plant of the second variety. In plants, pollen carries the male gametes (sperm) to the stigma, a sticky organ that traps pollen and allows the sperm to move down the pistil to the female gametes (ova) below. To prevent the pea plant that was receiving pollen from self-fertilizing and confounding his results, Mendel painstakingly removed all of the anthers

from the plant's flowers before they had a chance to mature.

Plants used in first-generation crosses were called **P₀**, or parental generation one ([\[link\]](#)). After each cross, Mendel collected the seeds belonging to the **P₀** plants and grew them the following season. These offspring were called the **F₁**, or the first filial (*filial* = offspring, daughter or son) generation. Once Mendel examined the characteristics in the **F₁** generation of plants, he allowed them to self-fertilize naturally. He then collected and grew the seeds from the **F₁** plants to produce the **F₂**, or second filial, generation. Mendel's experiments extended beyond the **F₂** generation to the **F₃** and **F₄** generations, and so on, but it was the ratio of characteristics in the **P₀ – F₁ – F₂** generations that were the most intriguing and became the basis for Mendel's postulates.

Cross between True-breeding Pea Plants

P generation



Violet
flowers



White
flowers

Hybridization of true-breeding plants



F₁ generation



All hybrid progeny have violet flowers

Self-fertilization of hybrid plants



F₂ generation



705
Violet flowers



224
White flowers

Garden Pea Characteristics Revealed the Basics of Heredity

In his 1865 publication, Mendel reported the results of his crosses involving seven different characteristics, each with two contrasting traits. A **trait** is defined as a variation in the physical appearance of a heritable characteristic. The characteristics included plant height, seed texture, seed color, flower color, pea pod size, pea pod color, and flower position. For the characteristic of flower color, for example, the two contrasting traits were white versus violet. To fully examine each characteristic, Mendel generated large numbers of F₁ and F₂ plants, reporting results from 19,959 F₂ plants alone. His findings were consistent.

What results did Mendel find in his crosses for flower color? First, Mendel confirmed that he had plants that bred true for white or violet flower color. Regardless of how many generations Mendel examined, all self-crossed offspring of parents with white flowers had white flowers, and all self-crossed offspring of parents with violet flowers had violet flowers. In addition, Mendel confirmed that, other than flower color, the pea plants were physically identical.

Once these validations were complete, Mendel applied the pollen from a plant with violet flowers to the stigma of a plant with white flowers. After

gathering and sowing the seeds that resulted from this cross, *Mendel found that 100 percent of the F₁ hybrid generation had violet flowers.* Conventional wisdom at that time (the blending theory) would have predicted the hybrid flowers to be pale violet or for hybrid plants to have equal numbers of white and violet flowers. In other words, the contrasting parental traits were expected to blend in the offspring. Instead, Mendel's results demonstrated that the white flower trait in the F₁ generation had completely disappeared.

Importantly, Mendel did not stop his experimentation there. He allowed the F₁ plants to self-fertilize and found that, of F₂-generation plants, 705 had violet flowers and 224 had white flowers. This was a ratio of 3.15 violet flowers per one white flower, or approximately 3:1. When Mendel transferred pollen from a plant with violet flowers to the stigma of a plant with white flowers and vice versa, he obtained about the same ratio regardless of which parent, male or female, contributed which trait. This is called a **reciprocal cross**—a paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross. For the other six characteristics Mendel examined, the F₁ and F₂ generations behaved in the same way as they had for flower color. One of the two traits would disappear completely from the F₁ generation only to reappear in the F₂ generation at a ratio of

approximately 3:1 ([link](#)).

The Results of Mendel's Garden Pea

Character	Hybridization	Contrasting P0 Traits	F1 Offspring Traits	F2 Offspring Traits	F2 Trait Ratios
Flower color	705 224	Violet vs. white	100 percent violet		3.15:1
Flower position	651 207	Axial vs. terminal	100 percent axial		3.14:1
Plant height	787 277	Tall vs. dwarf	100 percent tall		2.84:1
Seed texture	5,474 1,850	Round vs. wrinkled	100 percent round		2.96:1
Seed color	6,022 2,001	Yellow vs. green	100 percent yellow		3.01:1
Pea pod	882 314	Inflated vs. constricted	100 percent inflated		2.95:1

texture	299	vs.	constricted	percent
			constricted	inflated
Pea pod	428	Green	vs.	100
color	152	yellow	percent	green

2.82:1

Upon compiling his results for many thousands of plants, Mendel concluded that the characteristics could be divided into expressed and latent traits. He called these, respectively, dominant and recessive traits. **Dominant traits** are those that are inherited unchanged in a hybridization. **Recessive traits** become latent, or disappear, in the offspring of a hybridization. The recessive trait does, however, reappear in the progeny of the hybrid offspring. An example of a dominant trait is the violet-flower trait. For this same characteristic (flower color), white-colored flowers are a recessive trait. The fact that the recessive trait reappeared in the F₂ generation meant that the traits remained separate (not blended) in the plants of the F₁ generation. Mendel also proposed that plants possessed two copies of the trait for the flower-color characteristic, and that each parent transmitted one of its two copies to its offspring, where they came together. Moreover, the physical observation of a dominant trait could mean that the genetic composition of the organism included two dominant versions of the characteristic or that it included one dominant and one recessive version. Conversely, the observation of a recessive trait meant that the organism lacked any

dominant versions of this characteristic.

So why did Mendel repeatedly obtain 3:1 ratios in his crosses? To understand how Mendel deduced the basic mechanisms of inheritance that lead to such ratios, we must first review the laws of probability.

Probability Basics

Probabilities are mathematical measures of likelihood. The empirical probability of an event is calculated by dividing the number of times the event occurs by the total number of opportunities for the event to occur. It is also possible to calculate theoretical probabilities by dividing the number of times that an event is *expected* to occur by the number of times that it could occur. Empirical probabilities come from observations, like those of Mendel. Theoretical probabilities, on the other hand, come from knowing how the events are produced and assuming that the probabilities of individual outcomes are equal. A probability of one for some event indicates that it is guaranteed to occur, whereas a probability of zero indicates that it is guaranteed not to occur. An example of a genetic event is a round seed produced by a pea plant.

In one experiment, Mendel demonstrated that the probability of the event “round seed” occurring was one in the F₁ offspring of true-breeding parents, one

of which has round seeds and one of which has wrinkled seeds. When the F₁ plants were subsequently self-crossed, the probability of any given F₂ offspring having round seeds was now three out of four. In other words, in a large population of F₂ offspring chosen at random, 75 percent were expected to have round seeds, whereas 25 percent were expected to have wrinkled seeds. Using large numbers of crosses, Mendel was able to calculate probabilities and use these to predict the outcomes of other crosses.

The Product Rule and Sum Rule

Mendel demonstrated that pea plants transmit characteristics as discrete units from parent to offspring. As will be discussed, Mendel also determined that different characteristics, like seed color and seed texture, were transmitted independently of one another and could be considered in separate probability analyses. For instance, performing a cross between a plant with green, wrinkled seeds and a plant with yellow, round seeds still produced offspring that had a 3:1 ratio of green:yellow seeds (ignoring seed texture) and a 3:1 ratio of round:wrinkled seeds (ignoring seed color). The characteristics of color and texture did not influence each other.

The **product rule** of probability can be applied to this phenomenon of the independent transmission of

characteristics. The product rule states that the probability of two independent events occurring together can be calculated by multiplying the individual probabilities of each event occurring alone. To demonstrate the product rule, imagine that you are rolling a six-sided die (D) and flipping a penny (P) at the same time. The die may roll any number from 1–6 (D#), whereas the penny may turn up heads (PH) or tails (PT). The outcome of rolling the die has no effect on the outcome of flipping the penny and vice versa. There are 12 possible outcomes of this action ([\[link\]](#)), and each event is expected to occur with equal probability.

Twelve Equally Likely Outcomes of Rolling a Die and Flipping a Penny

Rolling Die	Flipping Penny
D1	PH
D1	PT
D2	PH
D2	PT
D3	PH
D3	PT
D4	PH
D4	PT

D5	Pr	
D5	Pr	
D6	Pr	
D6	Pr	

Of the 12 possible outcomes, the die has a 2/12 (or 1/6) probability of rolling a two, and the penny has a 6/12 (or 1/2) probability of coming up heads. By the product rule, the probability that you will obtain the combined outcome 2 and heads is: $(D_2) \times (P_H) = (1/6) \times (1/2)$ or 1/12 ([\[link\]](#)). Notice the word “and” in the description of the probability. The “and” is a signal to apply the product rule. For example, consider how the product rule is applied to the dihybrid cross: the probability of having both dominant traits in the F₂ progeny is the product of the probabilities of having the dominant trait for each characteristic, as shown here:

$$\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$$

On the other hand, the **sum rule** of probability is applied when considering two mutually exclusive outcomes that can come about by more than one pathway. The sum rule states that the probability of the occurrence of one event or the other event, of two mutually exclusive events, is the sum of their individual probabilities. Notice the word “or” in the description of the probability. The “or” indicates that you should apply the sum rule. In this case, let’s imagine you are flipping a penny (P) and a quarter (Q). What is the probability of one coin coming up

heads and one coin coming up tails? This outcome can be achieved by two cases: the penny may be heads (P_H) and the quarter may be tails (Q_T), or the quarter may be heads (Q_H) and the penny may be tails (P_T). Either case fulfills the outcome. By the sum rule, we calculate the probability of obtaining one head and one tail as $[(P_H) \times (Q_T)] + [(Q_H) \times (P_T)] = [(1/2) \times (1/2)] + [(1/2) \times (1/2)] = 1/2$ ([link]). You should also notice that we used the product rule to calculate the probability of P_H and Q_T , and also the probability of P_T and Q_H , before we summed them. Again, the sum rule can be applied to show the probability of having just one dominant trait in the F₂ generation of a dihybrid cross:

3 16 + 3 4 = 15 16

The Product Rule and Sum Rule

Product Rule

For independent events A and B, the probability (P) of them both occurring (A *and* B) is $(P_A \times P_B)$

Sum Rule

For mutually exclusive events A and B, the probability (P) that at least one occurs (A *or* B) is $(P_A + P_B)$

To use probability laws in practice, we must work with large sample sizes because small sample sizes are prone to deviations caused by chance. The large quantities of pea plants that Mendel examined allowed him calculate the probabilities of the traits appearing in his F₂ generation. As you will learn, this discovery meant that when parental traits were

known, the offspring's traits could be predicted accurately even before fertilization.

Section Summary

Working with garden pea plants, Mendel found that crosses between parents that differed by one trait produced F₁ offspring that all expressed the traits of one parent. Observable traits are referred to as dominant, and non-expressed traits are described as recessive. When the offspring in Mendel's experiment were self-crossed, the F₂ offspring exhibited the dominant trait or the recessive trait in a 3:1 ratio, confirming that the recessive trait had been transmitted faithfully from the original P₀ parent. Reciprocal crosses generated identical F₁ and F₂ offspring ratios. By examining sample sizes, Mendel showed that his crosses behaved reproducibly according to the laws of probability, and that the traits were inherited as independent events.

Two rules in probability can be used to find the expected proportions of offspring of different traits from different crosses. To find the probability of two or more independent events occurring together, apply the product rule and multiply the probabilities of the individual events. The use of the word "and" suggests the appropriate application of the product rule. To find the probability of two or more events

occurring in combination, apply the sum rule and add their individual probabilities together. The use of the word “or” suggests the appropriate application of the sum rule.

Review Questions

Mendel performed hybridizations by transferring pollen from the _____ of the male plant to the female ova.

1. anther
2. pistil
3. stigma
4. seed

A

Which is one of the seven characteristics that Mendel observed in pea plants?

1. flower size
 2. seed texture
 3. leaf shape
 4. stem color
-

B

Imagine you are performing a cross involving seed color in garden pea plants. What F₁ offspring would you expect if you cross true-breeding parents with green seeds and yellow seeds? Yellow seed color is dominant over green.

1. 100 percent yellow-green seeds
2. 100 percent yellow seeds
3. 50 percent yellow, 50 percent green seeds
4. 25 percent green, 75 percent yellow seeds

B

Consider a cross to investigate the pea pod texture trait, involving constricted or inflated pods. Mendel found that the traits behave according to a dominant/recessive pattern in which inflated pods were dominant. If you performed this cross and obtained 650 inflated-pod plants in the F₂ generation, approximately how many constricted-pod plants would you expect to have?

1. 600
2. 165
3. 217

C

A scientist pollinates a true-breeding pea plant with violet, terminal flowers with pollen from a true-breeding pea plant with white, axial flowers. Which of the following observations would most accurately describe the F₂ generation?

1. 75% violet flowers; 75% terminal flowers
2. 75% white flowers in a terminal position
3. 75% violet flowers; 75% axial flowers
4. 75% violet flowers in an axial position

C

Critical Thinking Questions

Describe one of the reasons why the garden pea was an excellent choice of model system for studying inheritance.

The garden pea is sessile and has flowers that

close tightly during self-pollination. These features help to prevent accidental or unintentional fertilizations that could have diminished the accuracy of Mendel's data.

How would you perform a reciprocal cross for the characteristic of stem height in the garden pea?

Two sets of P_0 parents would be used. In the first cross, pollen would be transferred from a true-breeding tall plant to the stigma of a true-breeding dwarf plant. In the second cross, pollen would be transferred from a true-breeding dwarf plant to the stigma of a true-breeding tall plant. For each cross, F_1 and F_2 offspring would be analyzed to determine if offspring traits were affected according to which parent donated each trait.

Mendel performs a cross using a true-breeding pea plant with round, yellow seeds and a true-breeding pea plant with green, wrinkled seeds. What is the probability that offspring will have green, round seeds? Calculate the probability for the F_1 and F_2 generations.

Since we are calculating the probability of two

independent events occurring simultaneously, we use the product rule.

F₁ generation: Since green seed color is recessive, there is a 0% probability that any plants in the F₁ generation will have green, round seeds.

F₂ generation: The probability of growing an F₂ generation plant with green seeds is $\frac{1}{4}$, while the probability of growing an F₂ generation plant with round seeds is $\frac{3}{4}$. We can use the product rule to then calculate the probability of a plant with green, round seeds:

$$\frac{1}{4} \times \frac{3}{4} = \frac{3}{16}$$

Calculate the probability of selecting a heart or a face card from a standard deck of cards. Is this outcome more or less likely than selecting a heart suit face card?

A standard deck of cards contains 52 cards, 13 of which are hearts and 12 of which are face cards.

Heart suit **or** face card: This calculation requires the sum rule since there are multiple pathways to successfully pulling a desired card.
 $\frac{13}{52} + \frac{12}{52} = \frac{25}{52} = 48\%$

The probability of selecting a heart suit or a face card is significantly more likely than the probability of selecting a heart suit face card ($3/52 = 5.8\%$).

Glossary

blending theory of inheritance

hypothetical inheritance pattern in which parental traits are blended together in the offspring to produce an intermediate physical appearance

continuous variation

inheritance pattern in which a character shows a range of trait values with small gradations rather than large gaps between them

discontinuous variation

inheritance pattern in which traits are distinct and are transmitted independently of one another

dominant

trait which confers the same physical appearance whether an individual has two copies of the trait or one copy of the dominant trait and one copy of the recessive trait

F₁

first filial generation in a cross; the offspring of the parental generation

F₂

second filial generation produced when F₁ individuals are self-crossed or fertilized with each other

hybridization

process of mating two individuals that differ with the goal of achieving a certain characteristic in their offspring

model system

species or biological system used to study a specific biological phenomenon to be applied to other different species

P₀

parental generation in a cross

product rule

probability of two independent events occurring simultaneously can be calculated by multiplying the individual probabilities of each event occurring alone

recessive

trait that appears “latent” or non-expressed when the individual also carries a dominant trait for that same characteristic; when

present as two identical copies, the recessive trait is expressed

reciprocal cross

paired cross in which the respective traits of the male and female in one cross become the respective traits of the female and male in the other cross

sum rule

probability of the occurrence of at least one of two mutually exclusive events is the sum of their individual probabilities

trait

variation in the physical appearance of a heritable characteristic

Characteristics and Traits

By the end of this section, you will be able to do the following:

- Explain the relationship between genotypes and phenotypes in dominant and recessive gene systems
- Develop a Punnett square to calculate the expected proportions of genotypes and phenotypes in a monohybrid cross
- Explain the purpose and methods of a test cross
- Identify non-Mendelian inheritance patterns such as incomplete dominance, codominance, recessive lethals, multiple alleles, and sex linkage

Physical characteristics are expressed through genes carried on chromosomes. The genetic makeup of peas consists of two similar, or homologous, copies of each chromosome, one from each parent. Each pair of homologous chromosomes has the same linear order of genes. In other words, peas are diploid organisms in that they have two copies of each chromosome. The same is true for many other plants and for virtually all animals. Diploid organisms produce haploid gametes, which contain one copy of each homologous chromosome that unite at fertilization to create a diploid zygote.

For cases in which a single gene controls a single characteristic, a diploid organism has two genetic

copies that may or may not encode the same version of that characteristic. Gene variants that arise by mutation and exist at the same relative locations on homologous chromosomes are called **alleles**.

Mendel examined the inheritance of genes with just two allele forms, but it is common to encounter more than two alleles for any given gene in a natural population.

Phenotypes and Genotypes

Two alleles for a given gene in a diploid organism are expressed and interact to produce physical characteristics. The observable traits expressed by an organism are referred to as its **phenotype**. An organism's underlying genetic makeup, consisting of both physically visible and non-expressed alleles, is called its **genotype**. Mendel's hybridization experiments demonstrate the difference between phenotype and genotype. When true-breeding plants in which one parent had yellow pods and one had green pods were cross-fertilized, all of the F₁ hybrid offspring had yellow pods. That is, the hybrid offspring were phenotypically identical to the true-breeding parent with yellow pods. However, we know that the allele donated by the parent with green pods was not simply lost because it reappeared in some of the F₂ offspring. Therefore, the F₁ plants must have been genotypically different from the parent with yellow pods.

The P₁ plants that Mendel used in his experiments were each homozygous for the trait he was studying. Diploid organisms that are **homozygous** at a given gene, or locus, have two identical alleles for that gene on their homologous chromosomes. Mendel's parental pea plants always bred true because both of the gametes produced carried the same trait. When P₁ plants with contrasting traits were cross-fertilized, all of the offspring were **heterozygous** for the contrasting trait, meaning that their genotype reflected that they had different alleles for the gene being examined.

Dominant and Recessive Alleles

Our discussion of homozygous and heterozygous organisms brings us to why the F₁ heterozygous offspring were identical to one of the parents, rather than expressing both alleles. In all seven pea-plant characteristics, one of the two contrasting alleles was dominant, and the other was recessive. Mendel called the dominant allele the expressed unit factor; the recessive allele was referred to as the latent unit factor. We now know that these so-called unit factors are actually genes on homologous chromosome pairs. For a gene that is expressed in a dominant and recessive pattern, homozygous dominant and heterozygous organisms will look identical (that is, they will have different genotypes but the same phenotype). The recessive allele will only be observed in homozygous recessive

individuals ([\[link\]](#)).

**Human Inheritance in
Dominant and Recessive
Patterns**

Dominant Traits

Achondroplasia

Brachydactyly

Huntington’s disease

Marfan syndrome

Neurofibromatosis

Widow’s peak

Wooly hair

Recessive Traits

Albinism

Cystic fibrosis

Duchenne muscular
dystrophy

Galaetosemia

Phenylketonuria

Sickle cell anemia

Tay-Sachs disease

Several conventions exist for referring to genes and alleles. For the purposes of this chapter, we will abbreviate genes using the first letter of the gene’s corresponding dominant trait. For example, violet is the dominant trait for a pea plant’s flower color, so the flower-color gene would be abbreviated as *V* (note that it is customary to italicize gene designations). Furthermore, we will use uppercase and lowercase letters to represent dominant and recessive alleles, respectively. Therefore, we would refer to the genotype of a homozygous dominant

pea plant with violet flowers as VV , a homozygous recessive pea plant with white flowers as vv , and a heterozygous pea plant with violet flowers as Vv .

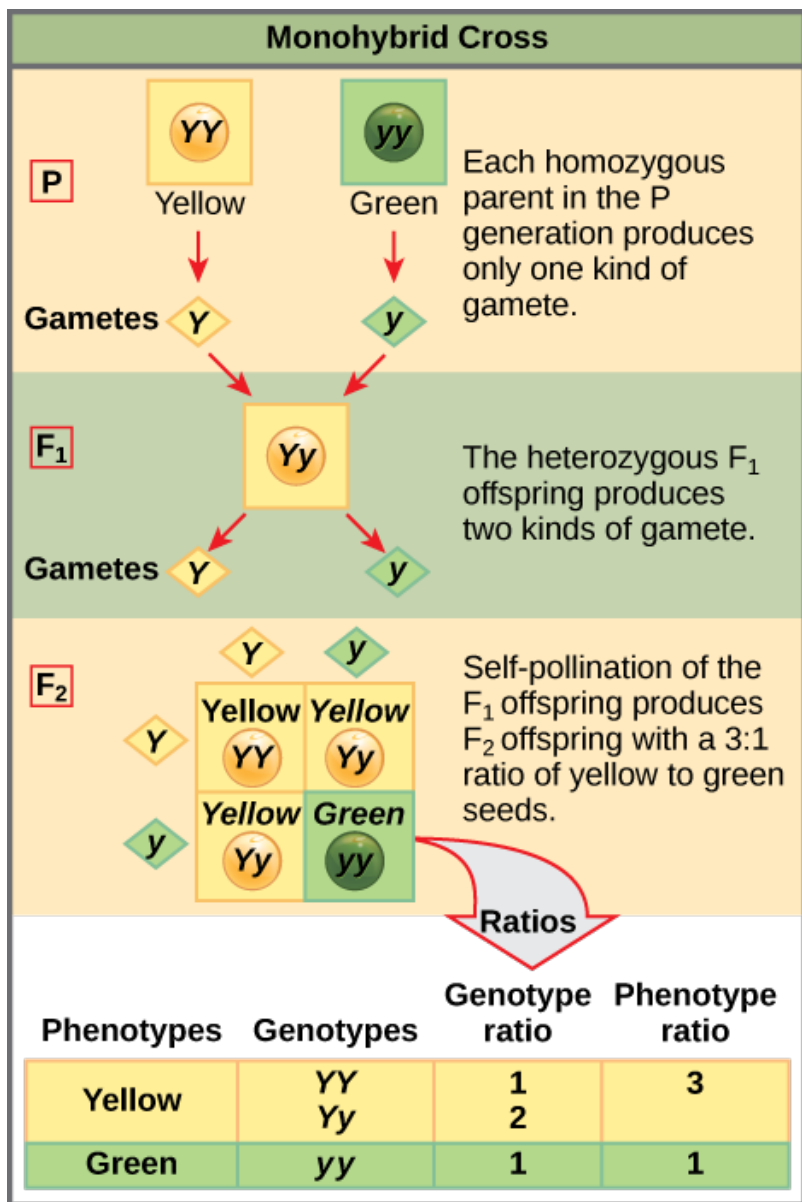
In the P generation, pea plants that are true-breeding for the dominant yellow phenotype are crossed with plants with the recessive green phenotype. This cross produces F_1 heterozygotes with a yellow phenotype. Punnett square analysis can be used to predict the genotypes of the F_2 generation.

The Punnett Square Approach for a Monohybrid Cross

When fertilization occurs between two true-breeding parents that differ in only one characteristic, the process is called a **monohybrid** cross, and the resulting offspring are monohybrids. Mendel performed seven monohybrid crosses involving contrasting traits for each characteristic. On the basis of his results in F_1 and F_2 generations, Mendel postulated that each parent in the monohybrid cross contributed one of two paired unit factors to each offspring, and every possible combination of unit factors was equally likely.

To demonstrate a monohybrid cross, consider the case of true-breeding pea plants with yellow versus green pea seeds. The dominant seed color is yellow; therefore, the parental genotypes were YY for the plants with yellow seeds and yy for the plants with

green seeds, respectively. A **Punnett square**, devised by the British geneticist Reginald Punnett, can be drawn that applies the rules of probability to predict the possible outcomes of a genetic cross or mating and their expected frequencies. To prepare a Punnett square, all possible combinations of the parental alleles are listed along the top (for one parent) and side (for the other parent) of a grid, representing their meiotic segregation into haploid gametes. Then the combinations of egg and sperm are made in the boxes in the table to show which alleles are combining. Each box then represents the diploid genotype of a zygote, or fertilized egg, that could result from this mating. Because each possibility is equally likely, genotypic ratios can be determined from a Punnett square. If the pattern of inheritance (dominant or recessive) is known, the phenotypic ratios can be inferred as well. For a monohybrid cross of two true-breeding parents, each parent contributes one type of allele. In this case, only one genotype is possible. All offspring are Yy and have yellow seeds ([\[link\]](#)).



A self-cross of one of the Yy heterozygous offspring can be represented in a 2×2 Punnett square because each parent can donate one of two different alleles. Therefore, the offspring can potentially have

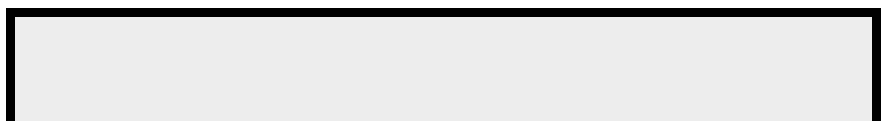
one of four allele combinations: YY , Yy , yY , or yy ([\[link\]](#)). Notice that there are two ways to obtain the Yy genotype: a Y from the egg and a y from the sperm, or a y from the egg and a Y from the sperm. Both of these possibilities must be counted. Recall that Mendel's pea-plant characteristics behaved in the same way in reciprocal crosses. Therefore, the two possible heterozygous combinations produce offspring that are genotypically and phenotypically identical despite their dominant and recessive alleles deriving from different parents. They are grouped together. Because fertilization is a random event, we expect each combination to be equally likely and for the offspring to exhibit a ratio of $YY:Yy:yy$ genotypes of 1:2:1 ([\[link\]](#)). Furthermore, because the YY and Yy offspring have yellow seeds and are phenotypically identical, applying the sum rule of probability, we expect the offspring to exhibit a phenotypic ratio of 3 yellow:1 green. Indeed, working with large sample sizes, Mendel observed approximately this ratio in every F_2 generation resulting from crosses for individual traits.

Mendel validated these results by performing an F_3 cross in which he self-crossed the dominant- and recessive-expressing F_2 plants. When he self-crossed the plants expressing green seeds, all of the offspring had green seeds, confirming that all green seeds had homozygous genotypes of yy . When he self-crossed the F_2 plants expressing yellow seeds, he

found that one-third of the plants bred true, and two-thirds of the plants segregated at a 3:1 ratio of yellow:green seeds. In this case, the true-breeding plants had homozygous (YY) genotypes, whereas the segregating plants corresponded to the heterozygous (Yy) genotype. When these plants self-fertilized, the outcome was just like the F₁ self-fertilizing cross.

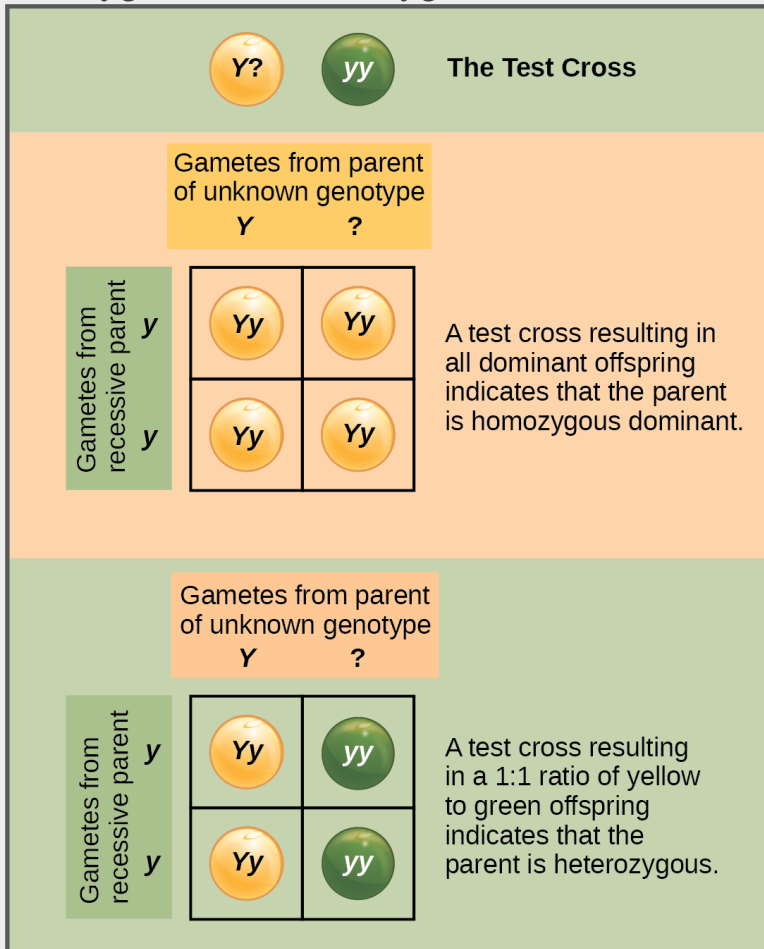
The Test Cross Distinguishes the Dominant Phenotype

Beyond predicting the offspring of a cross between known homozygous or heterozygous parents, Mendel also developed a way to determine whether an organism that expressed a dominant trait was a heterozygote or a homozygote. Called the **test cross**, this technique is still used by plant and animal breeders. In a test cross, the dominant-expressing organism is crossed with an organism that is homozygous recessive for the same characteristic. If the dominant-expressing organism is a homozygote, then all F₁ offspring will be heterozygotes expressing the dominant trait ([\[link\]](#)). Alternatively, if the dominant expressing organism is a heterozygote, the F₁ offspring will exhibit a 1:1 ratio of heterozygotes and recessive homozygotes ([\[link\]](#)). The test cross further validates Mendel's postulate that pairs of unit factors segregate equally.



Visual Connection

A test cross can be performed to determine whether an organism expressing a dominant trait is a homozygote or a heterozygote.



In pea plants, round peas (R) are dominant to wrinkled peas (r). You do a test cross between a pea plant with wrinkled peas (genotype rr) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the round pea

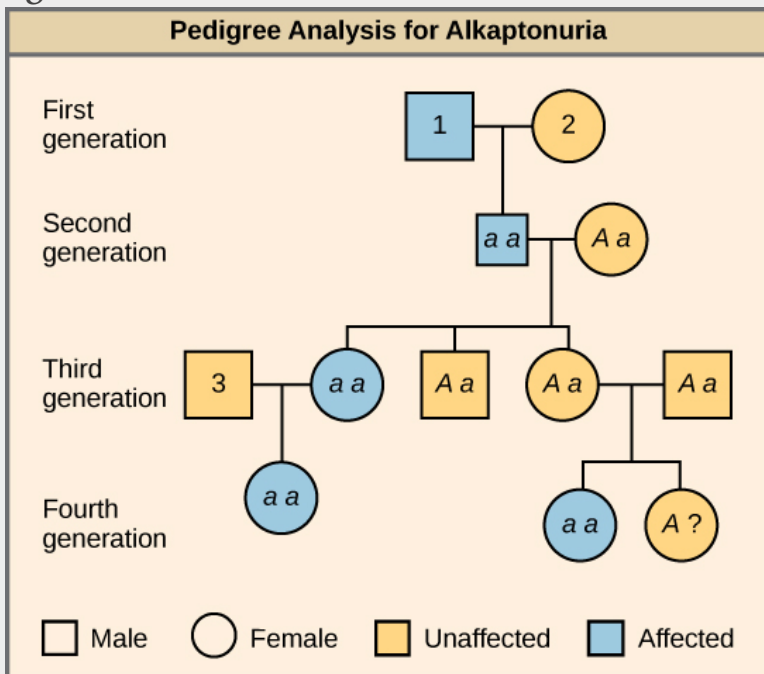
parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

Many human diseases are genetically inherited. A healthy person in a family in which some members suffer from a recessive genetic disorder may want to know if he or she has the disease-causing gene and what risk exists of passing the disorder on to his or her offspring. Of course, doing a test cross in humans is unethical and impractical. Instead, geneticists use **pedigree analysis** to study the inheritance pattern of human genetic diseases ([\[link\]](#)).

Visual Connection

Alkaptonuria is a recessive genetic disorder in which two amino acids, phenylalanine and tyrosine, are not properly metabolized. Affected individuals may have darkened skin and brown urine, and may suffer joint damage and other complications. In this pedigree, individuals with the disorder are indicated in blue and have the genotype *aa*. Unaffected individuals are indicated in yellow and have the genotype *AA* or *Aa*. Note

that it is often possible to determine a person's genotype from the genotype of their offspring. For example, if neither parent has the disorder but their child does, they must be heterozygous. Two individuals on the pedigree have an unaffected phenotype but unknown genotype. Because they do not have the disorder, they must have at least one normal allele, so their genotype gets the "A?" designation.



What are the genotypes of the individuals labeled 1, 2, and 3?

These pink flowers of a heterozygote snapdragon result from incomplete dominance. (credit:

“storebukkebruse”/Flickr) Four different alleles exist for the rabbit coat color (C) gene. As seen in comparing the wild-type *Drosophila* (left) and the *Antennapedia* mutant (right), the *Antennapedia* mutant has legs on its head in place of antennae. In *Drosophila*, several genes determine eye color. The genes for white and vermilion eye colors are located on the X chromosome. Others are located on the autosomes. Clockwise from top left are brown, cinnabar, sepia, vermilion, white, and red. Red eye color is wild-type and is dominant to white eye color. The son of a woman who is a carrier of a recessive X-linked disorder will have a 50 percent chance of being affected. A daughter will not be affected, but she will have a 50 percent chance of being a carrier like her mother. The neuron in the center of this micrograph (yellow) has nuclear inclusions characteristic of Huntington’s disease (orange area in the center of the neuron).

Huntington’s disease occurs when an abnormal dominant allele for the Huntington gene is present. (credit: Dr. Steven Finkbeiner, Gladstone Institute of Neurological Disease, The Taube-Koret Center for Huntington's Disease Research, and the University of California San Francisco/Wikimedia)

Alternatives to Dominance and Recessiveness

Mendel’s experiments with pea plants suggested that: (1) two “units” or alleles exist for every gene;

(2) alleles maintain their integrity in each generation (no blending); and (3) in the presence of the dominant allele, the recessive allele is hidden and makes no contribution to the phenotype. Therefore, recessive alleles can be “carried” and not expressed by individuals. Such heterozygous individuals are sometimes referred to as “carriers.” Further genetic studies in other plants and animals have shown that much more complexity exists, but that the fundamental principles of Mendelian genetics still hold true. In the sections to follow, we consider some of the extensions of Mendelism. If Mendel had chosen an experimental system that exhibited these genetic complexities, it’s possible that he would not have understood what his results meant.

Incomplete Dominance

Mendel’s results, that traits are inherited as dominant and recessive pairs, contradicted the view at that time that offspring exhibited a blend of their parents’ traits. However, the heterozygote phenotype occasionally does appear to be intermediate between the two parents. For example, in the snapdragon, *Antirrhinum majus* ([\[link\]](#)), a cross between a homozygous parent with white flowers (C^WC^W) and a homozygous parent with red flowers (C^RC^R) will produce offspring with pink flowers (C^RC^W). (Note that different genotypic abbreviations are used for Mendelian extensions to

distinguish these patterns from simple dominance and recessiveness.) This pattern of inheritance is described as **incomplete dominance**, denoting the expression of two contrasting alleles such that the individual displays an intermediate phenotype. The allele for red flowers is incompletely dominant over the allele for white flowers. However, the results of a heterozygote self-cross can still be predicted, just as with Mendelian dominant and recessive crosses. In this case, the genotypic ratio would be 1 $C_R C_R$:2 $C_R C_W$:1 $C_W C_W$, and the phenotypic ratio would be 1:2:1 for red:pink:white.



Codominance

A variation on incomplete dominance is **codominance**, in which both alleles for the same characteristic are simultaneously expressed in the heterozygote. An example of codominance is the





MN blood groups of humans. The M and N alleles are expressed in the form of an M or N antigen present on the surface of red blood cells.

Homozygotes ($LM LM$ and $LN LN$) express either the M or the N allele, and heterozygotes ($LM LN$) express both alleles equally. In a self-cross between heterozygotes expressing a codominant trait, the three possible offspring genotypes are phenotypically distinct. However, the 1:2:1 genotypic ratio characteristic of a Mendelian monohybrid cross still applies.

Multiple Alleles

Mendel implied that only two alleles, one dominant and one recessive, could exist for a given gene. We now know that this is an oversimplification. Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level such that many combinations of two alleles are observed. Note that when many alleles exist for the same gene, the convention is to denote the most common phenotype or genotype among wild animals as the **wild type** (often abbreviated “+”); this is considered the standard or norm. All other phenotypes or genotypes are considered **variants** of this standard, meaning that they deviate from the wild type. The variant may be recessive or dominant to the wild-type allele.

An example of multiple alleles is coat color in rabbits ([link](#)). Here, four alleles exist for the *c* gene. The wild-type version, $C+C+$, is expressed as brown fur. The chinchilla phenotype, $c^{ch}c^{ch}$, is expressed as black-tipped white fur. The Himalayan phenotype, $chch$, has black fur on the extremities and white fur elsewhere. Finally, the albino, or “colorless” phenotype, cc , is expressed as white fur. In cases of multiple alleles, dominance hierarchies can exist. In this case, the wild-type allele is dominant over all the others, chinchilla is incompletely dominant over Himalayan and albino, and Himalayan is dominant over albino. This hierarchy, or allelic series, was revealed by observing the phenotypes of each possible heterozygote offspring.

Allele			
C	c^{ch}	c^h	c
Genotype			
CC	$c^{ch}c^{ch}$	$c^h c^h$	cc
Phenotype			
WILD TYPE: Brown fur	CHINCHILLA: Black-tipped white fur	HIMALAYAN: White fur with black paws, nose, ears, tail	ALBINO: White fur
			

The complete dominance of a wild-type phenotype over all other mutants often occurs as an effect of

“dosage” of a specific gene product, such that the wild-type allele supplies the correct amount of gene product whereas the mutant alleles cannot. For the allelic series in rabbits, the wild-type allele may supply a given dosage of fur pigment, whereas the mutants supply a lesser dosage or none at all. Interestingly, the Himalayan phenotype is the result of an allele that produces a temperature-sensitive gene product that only produces pigment in the cooler extremities of the rabbit’s body.

Alternatively, one mutant allele can be dominant over all other phenotypes, including the wild type. This may occur when the mutant allele somehow interferes with the genetic message so that even a heterozygote with one wild-type allele copy expresses the mutant phenotype. One way in which the mutant allele can interfere is by enhancing the function of the wild-type gene product or changing its distribution in the body. One example of this is the *Antennapedia* mutation in *Drosophila* ([\[link\]](#)). In this case, the mutant allele expands the distribution of the gene product, and as a result, the *Antennapedia* heterozygote develops legs on its head where its antennae should be.



Evolution Connection

Multiple Alleles Confer Drug Resistance in the Malaria Parasite

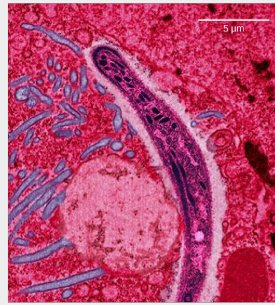
Malaria is a parasitic disease in humans that is transmitted by infected female mosquitoes, including *Anopheles gambiae* ([\[link\]a](#)), and is characterized by cyclic high fevers, chills, flu-like symptoms, and severe anemia. *Plasmodium falciparum* and *P. vivax* are the most common causative agents of malaria, and *P. falciparum* is the most deadly ([\[link\]b](#)). When promptly and correctly treated, *P. falciparum* malaria has a mortality rate of 0.1 percent. However, in some parts of the world, the parasite has evolved

resistance to commonly used malaria treatments, so the most effective malarial treatments can vary by geographic region.

The (a) *Anopheles gambiae*, or African malaria mosquito, acts as a vector in the transmission to humans of the malaria-causing parasite (b) *Plasmodium falciparum*, here visualized using false-color transmission electron microscopy. (credit a: James D. Gathany; credit b: Ute Frevert; false color by Margaret Shear; scale-bar data from Matt Russell)



(a)



(b)

In Southeast Asia, Africa, and South America, *P. falciparum* has developed resistance to the anti-malarial drugs chloroquine, mefloquine, and sulfadoxine-pyrimethamine. *P. falciparum*, which is haploid during the life stage in which it is infectious to humans, has evolved multiple drug-resistant mutant alleles of the *dhps* gene. Varying degrees of sulfadoxine resistance are associated with each of these alleles. Being haploid, *P. falciparum* needs only one drug-resistant allele to express this trait.

In Southeast Asia, different sulfadoxine-resistant alleles of the *dhps* gene are localized to different

geographic regions. This is a common evolutionary phenomenon that occurs because drug-resistant mutants arise in a population and interbreed with other *P. falciparum* isolates in close proximity. Sulfadoxine-resistant parasites cause considerable human hardship in regions where this drug is widely used as an over-the-counter malaria remedy. As is common with pathogens that multiply to large numbers within an infection cycle, *P. falciparum* evolves relatively rapidly (over a decade or so) in response to the selective pressure of commonly used anti-malarial drugs. For this reason, scientists must constantly work to develop new drugs or drug combinations to combat the worldwide malaria burden. [\[footnote\]](#)

Sumiti Vinayak, et al., “Origin and Evolution of Sulfadoxine Resistant *Plasmodium falciparum*,” *Public Library of Science Pathogens* 6, no. 3 (2010): e1000830, doi:10.1371/journal.ppat.1000830.

X-Linked Traits

In humans, as well as in many other animals and some plants, the sex of the individual is determined by sex chromosomes. The sex chromosomes are one pair of non-homologous chromosomes. Until now, we have only considered inheritance patterns among non-sex chromosomes, or **autosomes**. In addition to 22 homologous pairs of autosomes,

human females have a homologous pair of X chromosomes, whereas human males have an XY chromosome pair. Although the Y chromosome contains a small region of similarity to the X chromosome so that they can pair during meiosis, the Y chromosome is much shorter and contains many fewer genes. When a gene being examined is present on the X chromosome, but not on the Y chromosome, it is said to be **X-linked**.

Eye color in *Drosophila* was one of the first X-linked traits to be identified. Thomas Hunt Morgan mapped this trait to the X chromosome in 1910. Like humans, *Drosophila* males have an XY chromosome pair, and females are XX. In flies, the wild-type eye color is red (X^W) and it is dominant to white eye color (x^w) ([\[link\]](#)). Because of the location of the eye-color gene, reciprocal crosses do not produce the same offspring ratios. Males are said to be **hemizygous**, because they have only one allele for any X-linked characteristic. Hemizyosity makes the descriptions of dominance and recessiveness irrelevant for XY males. *Drosophila* males lack a second allele copy on the Y chromosome; that is, their genotype can only be X^WY or x^wY . In contrast, females have two allele copies of this gene and can be X^WX^W , X^Wx^w , or x^wx^w .



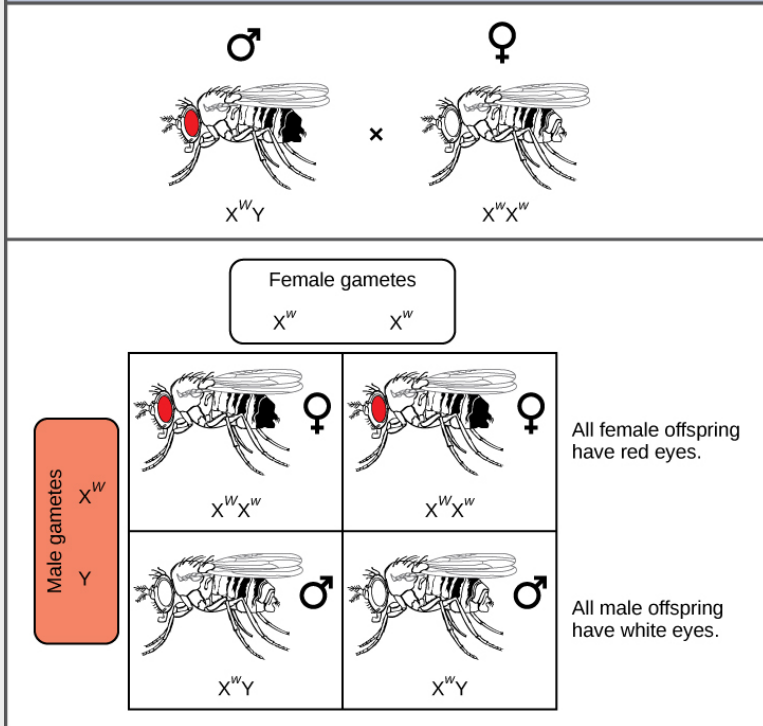
In an X-linked cross, the genotypes of F₁ and F₂ offspring depend on whether the recessive trait was expressed by the male or the female in the P₁ generation. With regard to *Drosophila* eye color, when the P₁ male expresses the white-eye phenotype and the female is homozygous red-eyed, all members of the F₁ generation exhibit red eyes ([\[link\]](#)). The F₁ females are heterozygous (X^wX^w), and the males are all X^wY , having received their X chromosome from the homozygous dominant P₁ female and their Y chromosome from the P₁ male. A subsequent cross between the X^wX^w female and the X^wY male would produce only red-eyed females (with X^wX^w or X^wX^w genotypes) and both red- and

white-eyed males (with X^wY or X^wY genotypes). Now, consider a cross between a homozygous white-eyed female and a male with red eyes. The F_1 generation would exhibit only heterozygous red-eyed females (X^WX^w) and only white-eyed males (X^wY). Half of the F_2 females would be red-eyed (X^WX^w) and half would be white-eyed (X^wX^w). Similarly, half of the F_2 males would be red-eyed (X^WY) and half would be white-eyed (X^wY).

Visual Connection

Punnett square analysis is used to determine the ratio of offspring from a cross between a red-eyed male fruit fly and a white-eyed female fruit fly.

Punnett Square Analysis of a Sex-linked Trait



What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

Discoveries in fruit fly genetics can be applied to human genetics. When a female parent is homozygous for a recessive X-linked trait, she will pass the trait on to 100 percent of her offspring. Her male offspring are, therefore, destined to express the trait, as they will inherit their father's Y chromosome. In humans, the alleles for certain conditions (some forms of color blindness,

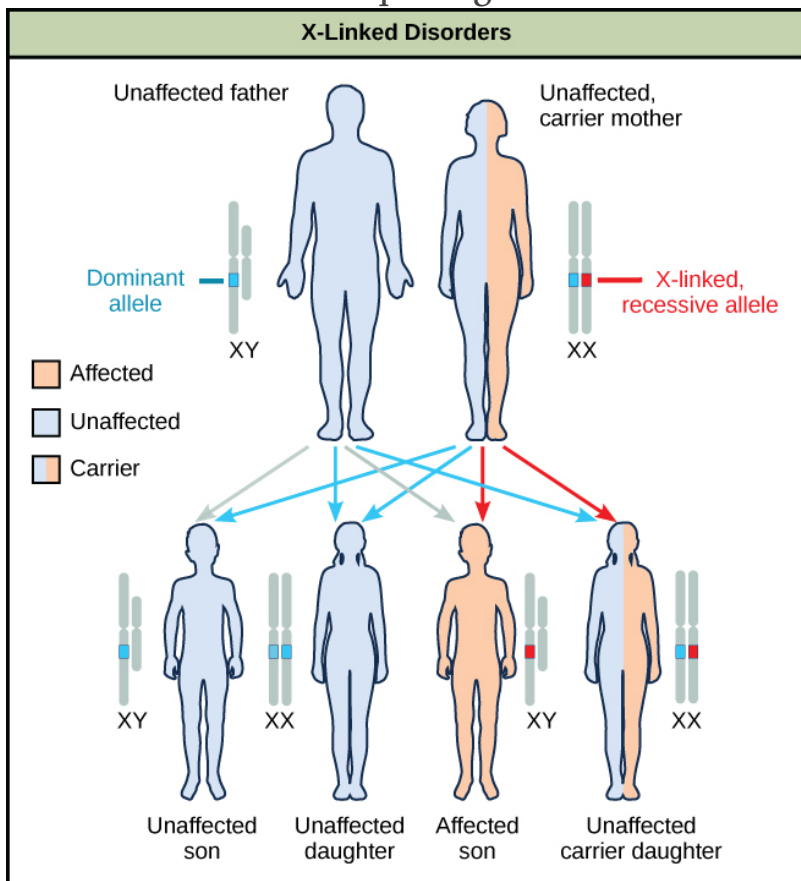
hemophilia, and muscular dystrophy) are X-linked. Females who are heterozygous for these diseases are said to be carriers and may not exhibit any phenotypic effects. These females will pass the disease to half of their sons and will pass carrier status to half of their daughters; therefore, recessive X-linked traits appear more frequently in males than females.

In some groups of organisms with sex chromosomes, the sex with the non-homologous sex chromosomes is the female rather than the male. This is the case for all birds. In this case, sex-linked traits will be more likely to appear in the female, in which they are hemizygous.

Human Sex-linked Disorders

Sex-linkage studies in Morgan's laboratory provided the fundamentals for understanding X-linked recessive disorders in humans, which include red-green color blindness, and Types A and B hemophilia. Because human males need to inherit only one recessive mutant X allele to be affected, X-linked disorders are disproportionately observed in males. Females must inherit recessive X-linked alleles from both of their parents in order to express the trait. When they inherit one recessive X-linked mutant allele and one dominant X-linked wild-type allele, they are carriers of the trait and are typically unaffected. Carrier females can manifest mild forms

of the trait due to the inactivation of the dominant allele located on one of the X chromosomes. However, female carriers can contribute the trait to their sons, resulting in the son exhibiting the trait, or they can contribute the recessive allele to their daughters, resulting in the daughters being carriers of the trait ([link]). Although some Y-linked recessive disorders exist, typically they are associated with infertility in males and are therefore not transmitted to subsequent generations.



Link to Learning

Watch this video to learn more about sex-linked traits.

https://www.openstax.org/l/sex-linked_trts

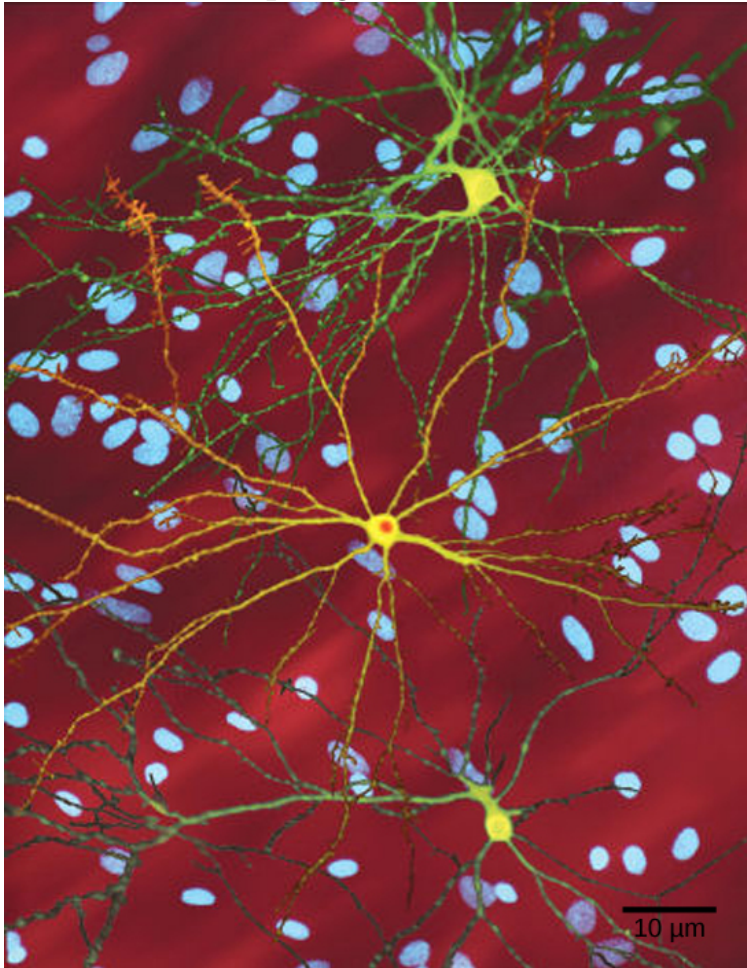
Lethality

A large proportion of genes in an individual's genome are essential for survival. Occasionally, a nonfunctional allele for an essential gene can arise by mutation and be transmitted in a population as long as individuals with this allele also have a wild-type, functional copy. The wild-type allele functions at a capacity sufficient to sustain life and is therefore considered to be dominant over the nonfunctional allele. However, consider two heterozygous parents that have a genotype of wild-type/nonfunctional mutant for a hypothetical essential gene. In one quarter of their offspring, we would expect to observe individuals that are homozygous recessive for the nonfunctional allele. Because the gene is essential, these individuals might fail to develop past fertilization, die *in utero*, or die later in life, depending on what life stage requires this gene. An inheritance pattern in which an allele is only lethal in the homozygous form and in which the heterozygote may be normal or have some altered nonlethal phenotype is referred to as **recessive lethal**.

For crosses between heterozygous individuals with a recessive lethal allele that causes death before birth when homozygous, only wild-type homozygotes and heterozygotes would be observed. The genotypic ratio would therefore be 2:1. In other instances, the recessive lethal allele might also exhibit a dominant (but not lethal) phenotype in the heterozygote. For instance, the recessive lethal *Curly* allele in *Drosophila* affects wing shape in the heterozygote form but is lethal in the homozygote.

A single copy of the wild-type allele is not always sufficient for normal functioning or even survival. The **dominant lethal** inheritance pattern is one in which an allele is lethal both in the homozygote and the heterozygote; this allele can only be transmitted if the lethality phenotype occurs after reproductive age. Individuals with mutations that result in dominant lethal alleles fail to survive even in the heterozygote form. Dominant lethal alleles are very rare because, as you might expect, the allele only lasts one generation and is not transmitted. However, just as the recessive lethal allele might not immediately manifest the phenotype of death, dominant lethal alleles also might not be expressed until adulthood. Once the individual reaches reproductive age, the allele may be unknowingly passed on, resulting in a delayed death in both generations. An example of this in humans is Huntington's disease, in which the nervous system gradually wastes away ([\[link\]](#)). People who are

heterozygous for the dominant Huntington allele (*Hh*) will inevitably develop the fatal disease. However, the onset of Huntington's disease may not occur until age 40, at which point the afflicted persons may have already passed the allele to 50 percent of their offspring.



Section Summary

When true-breeding or homozygous individuals that differ for a certain trait are crossed, all of the offspring will be heterozygotes for that trait. If the traits are inherited as dominant and recessive, the F₁ offspring will all exhibit the same phenotype as the parent homozygous for the dominant trait. If these heterozygous offspring are self-crossed, the resulting F₂ offspring will be equally likely to inherit gametes carrying the dominant or recessive trait, giving rise to offspring of which one quarter are homozygous dominant, half are heterozygous, and one quarter are homozygous recessive. Because homozygous dominant and heterozygous individuals are phenotypically identical, the observed traits in the F₂ offspring will exhibit a ratio of three dominant to one recessive.

Alleles do not always behave in dominant and recessive patterns. Incomplete dominance describes situations in which the heterozygote exhibits a phenotype that is intermediate between the homozygous phenotypes. Codominance describes the simultaneous expression of both of the alleles in the heterozygote. Although diploid organisms can only have two alleles for any given gene, it is common for more than two alleles of a gene to exist in a population. In humans, as in many animals and some plants, females have two X chromosomes and males have one X and one Y chromosome. Genes that are present on the X but not the Y chromosome are said to be X-linked, such that males only inherit

one allele for the gene, and females inherit two. Finally, some alleles can be lethal. Recessive lethal alleles are only lethal in homozygotes, but dominant lethal alleles are fatal in heterozygotes as well.

Visual Connection Questions

[\[link\]](#) In pea plants, round peas (R) are dominant to wrinkled peas (r). You do a test cross between a pea plant with wrinkled peas (genotype rr) and a plant of unknown genotype that has round peas. You end up with three plants, all which have round peas. From this data, can you tell if the round pea parent plant is homozygous dominant or heterozygous? If the round pea parent plant is heterozygous, what is the probability that a random sample of 3 progeny peas will all be round?

[\[link\]](#) You cannot be sure if the plant is homozygous or heterozygous as the data set is too small: by random chance, all three plants might have acquired only the dominant gene even if the recessive one is present. If the round pea parent is heterozygous, there is a one-eighth probability that a random sample of three progeny peas will all be round.

[\[link\]](#) What are the genotypes of the individuals labeled 1, 2, and 3?

[\[link\]](#) Individual 1 has the genotype aa . Individual 2 has the genotype Aa . Individual 3 has the genotype Aa .

[\[link\]](#) What ratio of offspring would result from a cross between a white-eyed male and a female that is heterozygous for red eye color?

[\[link\]](#) Half of the female offspring would be heterozygous ($X^W X^w$) with red eyes, and half would be homozygous recessive ($X^w X^w$) with white eyes. Half of the male offspring would be hemizygous dominant ($X^W Y$) with red eyes, and half would be hemizygous recessive ($X^w Y$) with white eyes.

Review Questions

The observable traits expressed by an organism are described as its _____.

1. phenotype

2. genotype
 3. alleles
 4. zygote
-

A

A recessive trait will be observed in individuals that are _____ for that trait.

1. heterozygous
 2. homozygous or heterozygous
 3. homozygous
 4. diploid
-

C

If black and white true-breeding mice are mated and the result is all gray offspring, what inheritance pattern would this be indicative of?

1. dominance
 2. codominance
 3. multiple alleles
 4. incomplete dominance
-

D

The ABO blood groups in humans are expressed as the I_A , I_B , and i alleles. The I_A allele encodes the A blood group antigen, I_B encodes B, and i encodes O. Both A and B are dominant to O. If a heterozygous blood type A parent ($I_A i$) and a heterozygous blood type B parent ($I_B i$) mate, one quarter of their offspring will have AB blood type ($I_A I_B$) in which both antigens are expressed equally. Therefore, ABO blood groups are an example of:

1. multiple alleles and incomplete dominance
2. codominance and incomplete dominance
3. incomplete dominance only
4. multiple alleles and codominance

D

In a mating between two individuals that are heterozygous for a recessive lethal allele that is expressed *in utero*, what genotypic ratio (homozygous dominant:heterozygous:homozygous recessive) would you expect to observe in the offspring?

1. 1:2:1
2. 3:1:1
3. 1:2:0
4. 0:2:1

C

If the allele encoding polydactyly (six fingers) is dominant why do most people have five fingers?

1. Genetic elements suppress the polydactyl gene.
2. Polydactyly is embryonic lethal.
3. The sixth finger is removed at birth.
4. The polydactyl allele is very rare in the human population.

D

A farmer raises black and white chickens. To his surprise, when the first generation of eggs hatch all the chickens are black with white speckles throughout their feathers. What should the farmer expect when the eggs laid after interbreeding the speckled chickens hatch?

1. All the offspring will be speckled.
2. 75% of the offspring will be speckled, and 25% will be black.
3. 50% of the offspring will be speckled, 25% will be black, and 25% will be white.
4. 50% of the offspring will be black and 50%

of the offspring will be white.

C

Critical Thinking Questions

The gene for flower position in pea plants exists as axial or terminal alleles. Given that axial is dominant to terminal, list all of the possible F₁ and F₂ genotypes and phenotypes from a cross involving parents that are homozygous for each trait. Express genotypes with conventional genetic abbreviations.

Because axial is dominant, the gene would be designated as *A*. F₁ would be all heterozygous *Aa* with axial phenotype. F₂ would have possible genotypes of *AA*, *Aa*, and *aa*; these would correspond to axial, axial, and terminal phenotypes, respectively.

Use a Punnett square to predict the offspring in a cross between a dwarf pea plant (homozygous recessive) and a tall pea plant (heterozygous). What is the phenotypic ratio of the offspring?

The Punnett square would be 2×2 and will have T and T along the top, and T and t along the left side. Clockwise from the top left, the genotypes listed within the boxes will be Tt , Tt , tt , and tt . The phenotypic ratio will be 1 tall:1 dwarf.

Can a human male be a carrier of red-green color blindness?

No, males can only express color blindness. They cannot carry it because an individual needs two X chromosomes to be a carrier.

Why is it more efficient to perform a test cross with a homozygous recessive donor than a homozygous dominant donor? How could the same information still be found with a homozygous dominant donor?

Using a homozygous recessive donor is more efficient because the genotype of the unknown parent can be determined in a single generation. If a homozygous dominant donor was used, the unknown genotype could still be determined. Instead of knowing the unknown genotype through the F_1 phenotype, the F_1

offspring would have to be self-crossed (as Mendel allowed his pea plants to self-pollinate) and the F₂ generation phenotypes would be used to determine the unknown F₀ genotype.

Glossary

allele

gene variations that arise by mutation and exist at the same relative locations on homologous chromosomes

autosomes

any of the non-sex chromosomes

codominance

in a heterozygote, complete and simultaneous expression of both alleles for the same characteristic

dominant lethal

inheritance pattern in which an allele is lethal both in the homozygote and the heterozygote; this allele can only be transmitted if the lethality phenotype occurs after reproductive age

genotype

underlying genetic makeup, consisting of both physically visible and non-expressed alleles, of an organism

hemizygous

presence of only one allele for a characteristic, as in X-linkage; hemizyosity makes descriptions of dominance and recessiveness irrelevant

heterozygous

having two different alleles for a given gene on the homologous chromosome

homozygous

having two identical alleles for a given gene on the homologous chromosome

incomplete dominance

in a heterozygote, expression of two contrasting alleles such that the individual displays an intermediate phenotype

monohybrid

result of a cross between two true-breeding parents that express different traits for only one characteristic

phenotype

observable traits expressed by an organism

Punnett square

visual representation of a cross between two individuals in which the gametes of each individual are denoted along the top and side of a grid, respectively, and the possible

zygotic genotypes are recombined at each box in the grid

recessive lethal

inheritance pattern in which an allele is only lethal in the homozygous form; the heterozygote may be normal or have some altered, nonlethal phenotype

sex-linked

any gene on a sex chromosome

test cross

cross between a dominant expressing individual with an unknown genotype and a homozygous recessive individual; the offspring phenotypes indicate whether the unknown parent is heterozygous or homozygous for the dominant trait

X-linked

gene present on the X, but not the Y chromosome

Laws of Inheritance

By the end of this section, you will be able to do the following:

- Explain Mendel's law of segregation and independent assortment in terms of genetics and the events of meiosis
- Use the forked-line method and the probability rules to calculate the probability of genotypes and phenotypes from multiple gene crosses
- Explain the effect of linkage and recombination on gamete genotypes
- Explain the phenotypic outcomes of epistatic effects between genes

Mendel generalized the results of his pea-plant experiments into four postulates, some of which are sometimes called “laws,” that describe the basis of dominant and recessive inheritance in diploid organisms. As you have learned, more complex extensions of Mendelism exist that do not exhibit the same F₂ phenotypic ratios (3:1). Nevertheless, these laws summarize the basics of classical genetics.

Pairs of Unit Factors, or Genes

Mendel proposed first that paired unit factors of heredity were transmitted faithfully from generation

to generation by the dissociation and reassociation of paired factors during gametogenesis and fertilization, respectively. After he crossed peas with contrasting traits and found that the recessive trait resurfaced in the F₂ generation, Mendel deduced that hereditary factors must be inherited as discrete units. This finding contradicted the belief at that time that parental traits were blended in the offspring.

The child in the photo expresses albinism, a recessive trait.

Alleles Can Be Dominant or Recessive

Mendel's **law of dominance** states that in a heterozygote, one trait will conceal the presence of another trait for the same characteristic. Rather than both alleles contributing to a phenotype, the dominant allele will be expressed exclusively. The recessive allele will remain "latent" but will be transmitted to offspring by the same manner in which the dominant allele is transmitted. The recessive trait will only be expressed by offspring that have two copies of this allele ([\[link\]](#)), and these offspring will breed true when self-crossed.

Since Mendel's experiments with pea plants, researchers have found that the law of dominance does not always hold true. Instead, several different patterns of inheritance have been found to exist.



Equal Segregation of Alleles

Observing that true-breeding pea plants with contrasting traits gave rise to F₁ generations that all expressed the dominant trait and F₂ generations that expressed the dominant and recessive traits in a 3:1 ratio, Mendel proposed the **law of segregation**. This law states that paired unit factors (genes) must segregate equally into gametes such that offspring have an equal likelihood of inheriting either factor.

For the F₂ generation of a monohybrid cross, the following three possible combinations of genotypes could result: homozygous dominant, heterozygous, or homozygous recessive. Because heterozygotes could arise from two different pathways (receiving one dominant and one recessive allele from either parent), and because heterozygotes and homozygous dominant individuals are phenotypically identical, the law supports Mendel's observed 3:1 phenotypic ratio. The equal segregation of alleles is the reason we can apply the Punnett square to accurately predict the offspring of parents with known genotypes. The physical basis of Mendel's law of segregation is the first division of meiosis, in which the homologous chromosomes with their different versions of each gene are segregated into daughter nuclei. The role of the meiotic segregation of chromosomes in sexual reproduction was not understood by the scientific community during Mendel's lifetime.

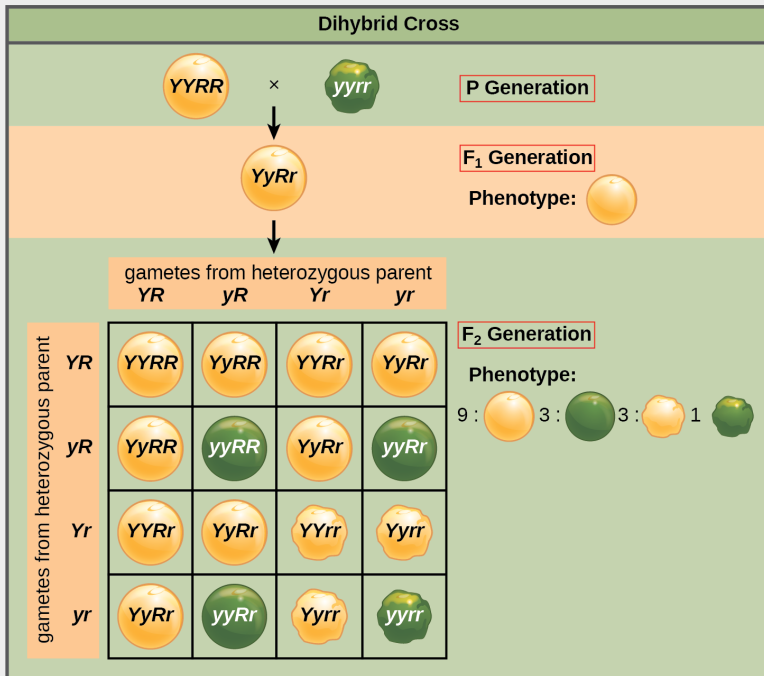
The forked-line method can be used to analyze a trihybrid cross. Here, the probability for color in the F₂ generation occupies the top row (3 yellow:1 green). The probability for shape occupies the second row (3 round: 1 wrinkled), and the probability for height occupies the third row (3 tall:1 dwarf). The probability for each possible combination of traits is calculated by multiplying the probability for each individual trait. Thus, the probability of F₂ offspring having yellow, round, and tall traits is $3 \times 3 \times 3$, or 27.

Independent Assortment

Mendel's **law of independent assortment** states that genes do not influence each other with regard to the sorting of alleles into gametes, and every possible combination of alleles for every gene is equally likely to occur. The independent assortment of genes can be illustrated by the **dihybrid** cross, a cross between two true-breeding parents that express different traits for two characteristics. Consider the characteristics of seed color and seed texture for two pea plants, one that has green, wrinkled seeds ($yyrr$) and another that has yellow, round seeds ($YYRR$). Because each parent is homozygous, the law of segregation indicates that the gametes for the green/wrinkled plant all are yr , and the gametes for the yellow/round plant are all YR . Therefore, the F_1 generation of offspring all are $YyRr$ ([\[link\]](#)).

Visual Connection

This dihybrid cross of pea plants involves the genes for seed color and texture.



In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between $PpYY$ and $ppYy$ pea plants? How many squares do you need to do a Punnett square analysis of this cross?

For the F₂ generation, the law of segregation requires that each gamete receive either an R allele or an r allele along with either a Y allele or a y allele. The law of independent assortment states that a gamete into which an r allele sorted would be equally likely to contain either a Y allele or a y allele. Thus, there are four equally likely gametes

that can be formed when the $YyRr$ heterozygote is self-crossed, as follows: YR , Yr , yR , and yr .

Arranging these gametes along the top and left of a 4×4 Punnett square ([\[link\]](#)) gives us 16 equally likely genotypic combinations. From these genotypes, we infer a phenotypic ratio of 9 round/yellow:3 round/green:3 wrinkled/yellow:1 wrinkled/green ([\[link\]](#)). These are the offspring ratios we would expect, assuming we performed the crosses with a large enough sample size.

Because of independent assortment and dominance, the 9:3:3:1 dihybrid phenotypic ratio can be collapsed into two 3:1 ratios, characteristic of any monohybrid cross that follows a dominant and recessive pattern. Ignoring seed color and considering only seed texture in the above dihybrid cross, we would expect that three quarters of the F_2 generation offspring would be round, and one quarter would be wrinkled. Similarly, isolating only seed color, we would assume that three quarters of the F_2 offspring would be yellow and one quarter would be green. The sorting of alleles for texture and color are independent events, so we can apply the product rule. Therefore, the proportion of round and yellow F_2 offspring is expected to be $(3/4) \times (3/4) = 9/16$, and the proportion of wrinkled and green offspring is expected to be $(1/4) \times (1/4) = 1/16$. These proportions are identical to those obtained using a Punnett square. Round, green and wrinkled, yellow offspring can also be calculated

using the product rule, as each of these genotypes includes one dominant and one recessive phenotype. Therefore, the proportion of each is calculated as $(3/4) \times (1/4) = 3/16$.

The law of independent assortment also indicates that a cross between yellow, wrinkled ($YYrr$) and green, round ($yyRR$) parents would yield the same F₁ and F₂ offspring as in the $YYRR \times yyrr$ cross.

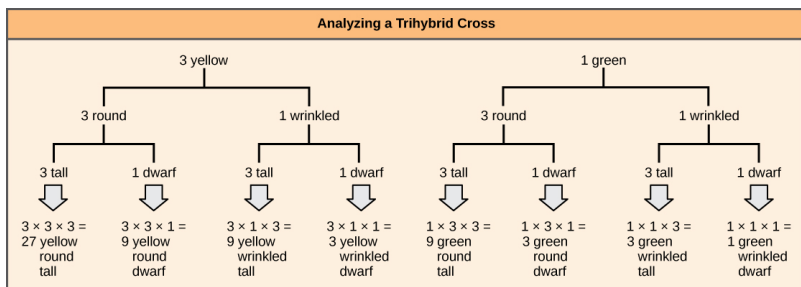
The physical basis for the law of independent assortment also lies in meiosis I, in which the different homologous pairs line up in random orientations. Each gamete can contain any combination of paternal and maternal chromosomes (and therefore the genes on them) because the orientation of tetrads on the metaphase plane is random.

Forked-Line Method

When more than two genes are being considered, the Punnett-square method becomes unwieldy. For instance, examining a cross involving four genes would require a 16×16 grid containing 256 boxes. It would be extremely cumbersome to manually enter each genotype. For more complex crosses, the forked-line and probability methods are preferred.

To prepare a forked-line diagram for a cross between F₁ heterozygotes resulting from a cross

between *AABBCC* and *aabbcc* parents, we first create rows equal to the number of genes being considered, and then segregate the alleles in each row on forked lines according to the probabilities for individual monohybrid crosses ([\[link\]](#)). We then multiply the values along each forked path to obtain the F₂ offspring probabilities. Note that this process is a diagrammatic version of the product rule. The values along each forked pathway can be multiplied because each gene assort independently. For a trihybrid cross, the F₂ phenotypic ratio is 27:9:9:9:3:3:3:1.



Probability Method

While the forked-line method is a diagrammatic approach to keeping track of probabilities in a cross, the probability method gives the proportions of offspring expected to exhibit each phenotype (or genotype) without the added visual assistance. Both methods make use of the product rule and consider the alleles for each gene separately. Earlier, we examined the phenotypic proportions for a trihybrid cross using the forked-line method; now we will use the probability method to examine the genotypic

proportions for a cross with even more genes.

For a trihybrid cross, writing out the forked-line method is tedious, albeit not as tedious as using the Punnett-square method. To fully demonstrate the power of the probability method, however, we can consider specific genetic calculations. For instance, for a tetrahybrid cross between individuals that are heterozygotes for all four genes, and in which all four genes are sorting independently and in a dominant and recessive pattern, what proportion of the offspring will be expected to be homozygous recessive for all four alleles? Rather than writing out every possible genotype, we can use the probability method. We know that for each gene, the fraction of homozygous recessive offspring will be $1/4$. Therefore, multiplying this fraction for each of the four genes, $(1/4) \times (1/4) \times (1/4) \times (1/4)$, we determine that $1/256$ of the offspring will be quadruply homozygous recessive.

For the same tetrahybrid cross, what is the expected proportion of offspring that have the dominant phenotype at all four loci? We can answer this question using phenotypic proportions, but let's do it the hard way—using genotypic proportions. The question asks for the proportion of offspring that are 1) homozygous dominant at *A* or heterozygous at *A*, and 2) homozygous at *B* or heterozygous at *B*, and so on. Noting the “or” and “and” in each circumstance makes clear where to apply the sum

and product rules. The probability of a homozygous dominant at *A* is $1/4$ and the probability of a heterozygote at *A* is $1/2$. The probability of the homozygote or the heterozygote is $1/4 + 1/2 = 3/4$ using the sum rule. The same probability can be obtained in the same way for each of the other genes, so that the probability of a dominant phenotype at *A* and *B* and *C* and *D* is, using the product rule, equal to $3/4 \times 3/4 \times 3/4 \times 3/4$, or $81/256$. If you are ever unsure about how to combine probabilities, returning to the forked-line method should make it clear.

Rules for Multihybrid Fertilization

Predicting the genotypes and phenotypes of offspring from given crosses is the best way to test your knowledge of Mendelian genetics. Given a multihybrid cross that obeys independent assortment and follows a dominant and recessive pattern, several generalized rules exist; you can use these rules to check your results as you work through genetics calculations ([\[link\]](#)). To apply these rules, first you must determine n , the number of heterozygous gene pairs (the number of genes segregating two alleles each). For example, a cross between *AaBb* and *AaBb* heterozygotes has an n of 2. In contrast, a cross between *AABb* and *AABb* has an n of 1 because *A* is not heterozygous.

General Rules for Multihybrid Crosses	
General Rule	Number of Heterozygous Gene Pairs
Number of different F1 gametes	2^n
Number of different F2 genotypes	3^n
Given dominant and recessive inheritance, the number of different F2 phenotypes	2^n

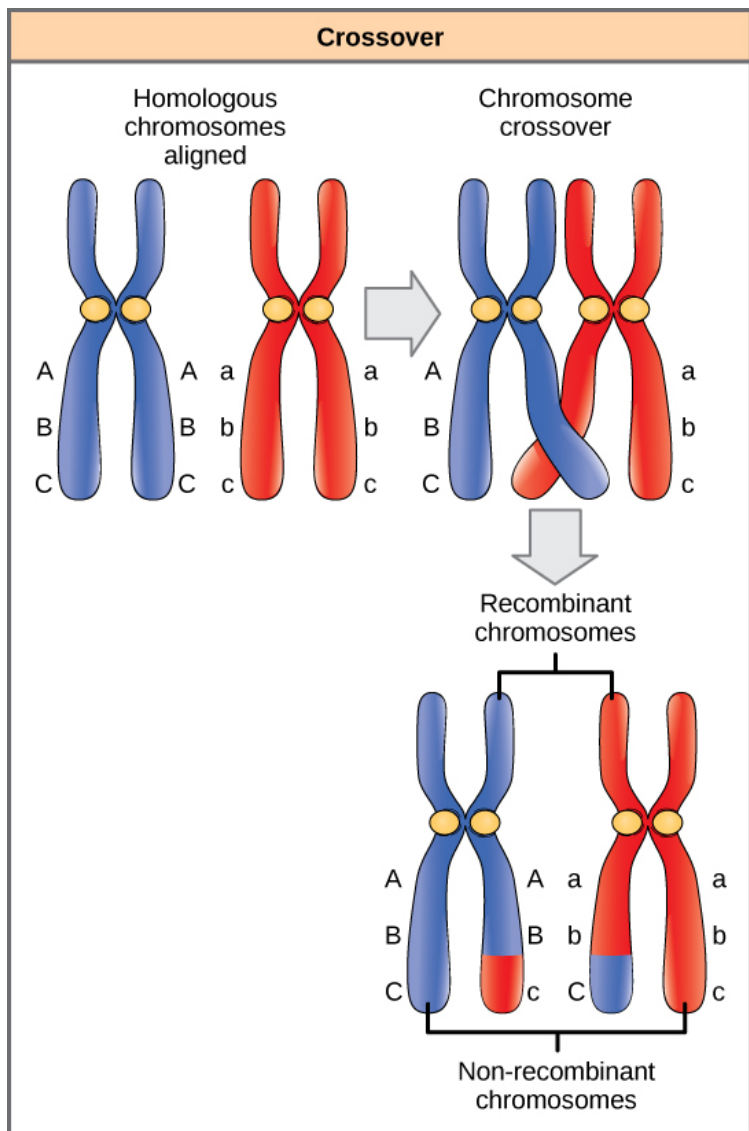
The process of crossover, or recombination, occurs when two homologous chromosomes align during meiosis and exchange a segment of genetic material. Here, the alleles for gene C were exchanged. The result is two recombinant and two non-recombinant chromosomes.

Linked Genes Violate the Law of Independent Assortment

Although all of Mendel's pea characteristics behaved according to the law of independent assortment, we now know that some allele combinations are not inherited independently of each other. Genes that are located on separate non-homologous chromosomes will always sort independently.

However, each chromosome contains hundreds or thousands of genes, organized linearly on chromosomes like beads on a string. The segregation of alleles into gametes can be influenced by **linkage**, in which genes that are located physically close to each other on the same chromosome are more likely to be inherited as a pair. However, because of the process of recombination, or “crossover,” it is possible for two genes on the same chromosome to behave independently, or as if they are not linked. To understand this, let’s consider the biological basis of gene linkage and recombination.

Homologous chromosomes possess the same genes in the same linear order. The alleles may differ on homologous chromosome pairs, but the genes to which they correspond do not. In preparation for the first division of meiosis, homologous chromosomes replicate and synapse. Like genes on the homologs align with each other. At this stage, segments of homologous chromosomes exchange linear segments of genetic material ([\[link\]](#)). This process is called *recombination*, or crossover, and it is a common genetic process. Because the genes are aligned during recombination, the gene order is not altered. Instead, the result of recombination is that maternal and paternal alleles are combined onto the same chromosome. Across a given chromosome, several recombination events may occur, causing extensive shuffling of alleles.



When two genes are located in close proximity on the same chromosome, they are considered linked, and their alleles tend to be transmitted through meiosis together. To exemplify this, imagine a dihybrid cross involving flower color and plant

height in which the genes are next to each other on the chromosome. If one homologous chromosome has alleles for tall plants and red flowers, and the other chromosome has genes for short plants and yellow flowers, then when the gametes are formed, the tall and red alleles will go together into a gamete and the short and yellow alleles will go into other gametes. These are called the parental genotypes because they have been inherited intact from the parents of the individual producing gametes. But unlike if the genes were on different chromosomes, there will be no gametes with tall and yellow alleles and no gametes with short and red alleles. If you create the Punnett square with these gametes, you will see that the classical Mendelian prediction of a 9:3:3:1 outcome of a dihybrid cross would not apply. As the distance between two genes increases, the probability of one or more crossovers between them increases, and the genes behave more like they are on separate chromosomes. Geneticists have used the proportion of recombinant gametes (the ones not like the parents) as a measure of how far apart genes are on a chromosome. Using this information, they have constructed elaborate maps of genes on chromosomes for well-studied organisms, including humans.

Mendel's seminal publication makes no mention of linkage, and many researchers have questioned whether he encountered linkage but chose not to

publish those crosses out of concern that they would invalidate his independent assortment postulate. The garden pea has seven chromosomes, and some have suggested that his choice of seven characteristics was not a coincidence. However, even if the genes he examined were not located on separate chromosomes, it is possible that he simply did not observe linkage because of the extensive shuffling effects of recombination.

Scientific Method Connection

Testing the Hypothesis of Independent Assortment

To better appreciate the amount of labor and ingenuity that went into Mendel's experiments, proceed through one of Mendel's dihybrid crosses.

Question: What will be the offspring of a dihybrid cross?

Background: Consider that pea plants mature in one growing season, and you have access to a large garden in which you can cultivate thousands of pea plants. There are several true-breeding plants with the following pairs of traits: tall plants with inflated pods, and dwarf plants with constricted pods. Before the plants have matured, you remove the pollen-producing organs from the tall/inflated plants in your crosses to prevent self-fertilization. Upon plant maturation, the plants are manually crossed by transferring pollen from the dwarf/

constricted plants to the stigmata of the tall/inflated plants.

Hypothesis: Both trait pairs will sort independently according to Mendelian laws. When the true-breeding parents are crossed, all of the F₁ offspring are tall and have inflated pods, which indicates that the tall and inflated traits are dominant over the dwarf and constricted traits, respectively. A self-cross of the F₁ heterozygotes results in 2,000 F₂ progeny.

Test the hypothesis: Because each trait pair sorts independently, the ratios of tall:dwarf and inflated:constricted are each expected to be 3:1.

The tall/dwarf trait pair is called T/t , and the inflated/constricted trait pair is designated I/i .

Each member of the F₁ generation therefore has a genotype of $TtIi$. Construct a grid analogous to [\[link\]](#), in which you cross two $TtIi$ individuals.

Each individual can donate four combinations of two traits: TI , Ti , tI , or ti , meaning that there are 16 possibilities of offspring genotypes. Because the T and I alleles are dominant, any individual having one or two of those alleles will express the tall or inflated phenotypes, respectively, regardless if they also have a t or i allele. Only individuals that are tt or ii will express the dwarf and constricted alleles, respectively. As shown in [\[link\]](#), you predict that you will observe the following offspring proportions: tall/inflated:tall/constricted:dwarf/inflated:dwarf/constricted in a 9:3:3:1 ratio. Notice from the grid that when considering the tall/dwarf

and inflated/constricted trait pairs in isolation, they are each inherited in 3:1 ratios.

This figure shows all possible combinations of offspring resulting from a dihybrid cross of pea plants that are heterozygous for the tall/dwarf and inflated/constricted alleles.

		<i>TtIi</i>			
		<i>TI</i>	<i>Ti</i>	<i>tI</i>	<i>ti</i>
<i>TtIi</i>	<i>TI</i>	<i>TTII</i>	<i>TTIi</i>	<i>TtII</i>	<i>TtIi</i>
	<i>Ti</i>	<i>TTIi</i>	<i>TTii</i>	<i>TtIi</i>	<i>Ttii</i>
	<i>tI</i>	<i>TtII</i>	<i>TtIi</i>	<i>ttII</i>	<i>ttIi</i>
	<i>ti</i>	<i>TtIi</i>	<i>Ttii</i>	<i>ttIi</i>	<i>ttii</i>

Test the hypothesis: You cross the dwarf and tall plants and then self-cross the offspring. For best results, this is repeated with hundreds or even thousands of pea plants. What special precautions should be taken in the crosses and in growing the plants?

Analyze your data: You observe the following plant phenotypes in the F₂ generation: 2706 tall/inflated, 930 tall/constricted, 888 dwarf/inflated,

and 300 dwarf/constricted. Reduce these findings to a ratio and determine if they are consistent with Mendelian laws.

Form a conclusion: Were the results close to the expected 9:3:3:1 phenotypic ratio? Do the results support the prediction? What might be observed if far fewer plants were used, given that alleles segregate randomly into gametes? Try to imagine growing that many pea plants, and consider the potential for experimental error. For instance, what would happen if it was extremely windy one day?

In mice, the mottled agouti coat color (A) is dominant to a solid coloration, such as black or gray. A gene at a separate locus (C) is responsible for pigment production. The recessive c allele does not produce pigment, and a mouse with the homozygous recessive cc genotype is albino regardless of the allele present at the A locus. Thus, the C gene is epistatic to the A gene.

Epistasis

Mendel's studies in pea plants implied that the sum of an individual's phenotype was controlled by genes (or as he called them, unit factors), such that every characteristic was distinctly and completely controlled by a single gene. In fact, single observable characteristics are almost always under

the influence of multiple genes (each with two or more alleles) acting in unison. For example, at least eight genes contribute to eye color in humans.

Link to Learning

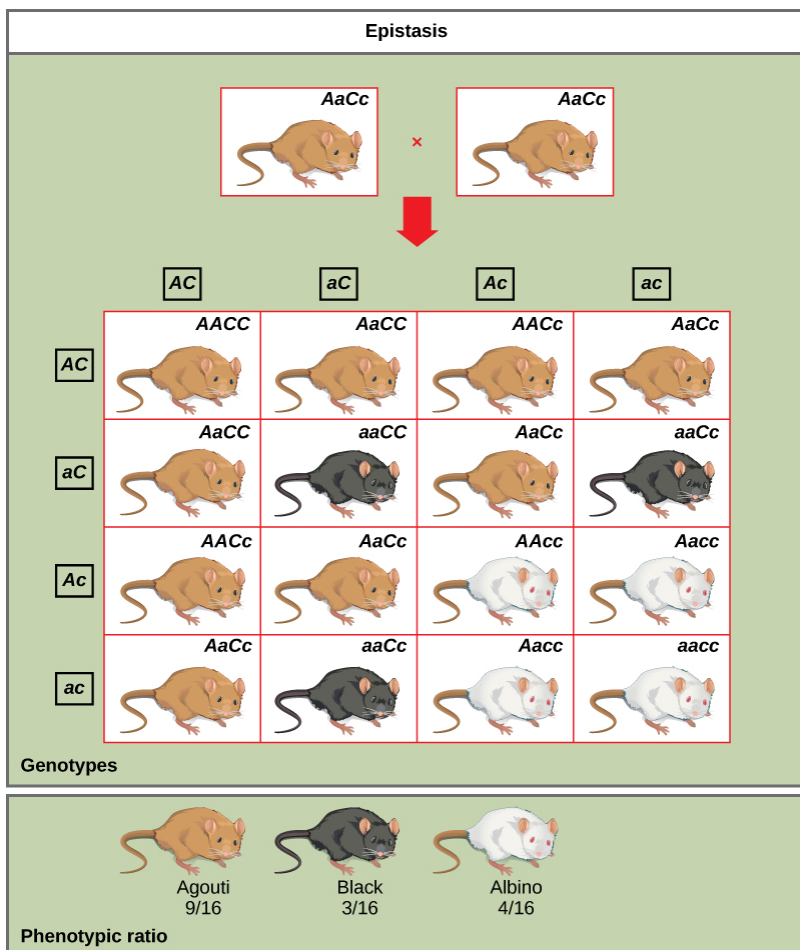
Eye color in humans is determined by multiple genes. Use the [Eye Color Calculator](#) to predict the eye color of children from parental eye color.

In some cases, several genes can contribute to aspects of a common phenotype without their gene products ever directly interacting. In the case of organ development, for instance, genes may be expressed sequentially, with each gene adding to the complexity and specificity of the organ. Genes may function in complementary or synergistic fashions, such that two or more genes need to be expressed simultaneously to affect a phenotype. Genes may also oppose each other, with one gene modifying the expression of another.

In **epistasis**, the interaction between genes is antagonistic, such that one gene masks or interferes with the expression of another. “Epistasis” is a word composed of Greek roots that mean “standing upon.” The alleles that are being masked or silenced are said to be hypostatic to the epistatic alleles that

are doing the masking. Often the biochemical basis of epistasis is a gene pathway in which the expression of one gene is dependent on the function of a gene that precedes or follows it in the pathway.

An example of epistasis is pigmentation in mice. The wild-type coat color, agouti (AA), is dominant to solid-colored fur (aa). However, a separate gene (C) is necessary for pigment production. A mouse with a recessive c allele at this locus is unable to produce pigment and is albino regardless of the allele present at locus A ([\[link\]](#)). Therefore, the genotypes $AAcc$, $Aacc$, and $aacc$ all produce the same albino phenotype. A cross between heterozygotes for both genes ($AaCc \times AaCc$) would generate offspring with a phenotypic ratio of 9 agouti:3 solid color:4 albino ([\[link\]](#)). In this case, the C gene is epistatic to the A gene.



Epistasis can also occur when a dominant allele masks expression at a separate gene. Fruit color in summer squash is expressed in this way.

Homozygous recessive expression of the W gene (ww) coupled with homozygous dominant or heterozygous expression of the Y gene (YY or Yy) generates yellow fruit, and the $wwyy$ genotype produces green fruit. However, if a dominant copy of the W gene is present in the homozygous or

heterozygous form, the summer squash will produce white fruit regardless of the Y alleles. A cross between white heterozygotes for both genes ($WwYy \times WwYy$) would produce offspring with a phenotypic ratio of 12 white:3 yellow:1 green.

Finally, epistasis can be reciprocal such that either gene, when present in the dominant (or recessive) form, expresses the same phenotype. In the shepherd's purse plant (*Capsella bursa-pastoris*), the characteristic of seed shape is controlled by two genes in a dominant epistatic relationship. When the genes A and B are both homozygous recessive ($aabb$), the seeds are ovoid. If the dominant allele for either of these genes is present, the result is triangular seeds. That is, every possible genotype other than $aabb$ results in triangular seeds, and a cross between heterozygotes for both genes ($AaBb \times AaBb$) would yield offspring with a phenotypic ratio of 15 triangular:1 ovoid.

As you work through genetics problems, keep in mind that any single characteristic that results in a phenotypic ratio that totals 16 is typical of a two-gene interaction. Recall the phenotypic inheritance pattern for Mendel's dihybrid cross, which considered two noninteracting genes—9:3:3:1. Similarly, we would expect interacting gene pairs to also exhibit ratios expressed as 16 parts. Note that we are assuming the interacting genes are not linked; they are still assorting independently into

gametes.

Link to Learning

For an excellent review of Mendel's experiments and to perform your own crosses and identify patterns of inheritance, visit the [Mendel's Peas](#) web lab.

Section Summary

Mendel postulated that genes (characteristics) are inherited as pairs of alleles (traits) that behave in a dominant and recessive pattern. Alleles segregate into gametes such that each gamete is equally likely to receive either one of the two alleles present in a diploid individual. In addition, genes are assorted into gametes independently of one another. That is, alleles are generally not more likely to segregate into a gamete with a particular allele of another gene. A dihybrid cross demonstrates independent assortment when the genes in question are on different chromosomes or distant from each other on the same chromosome. For crosses involving more than two genes, use the forked line or probability methods to predict offspring genotypes

and phenotypes rather than a Punnett square.

Although chromosomes sort independently into gametes during meiosis, Mendel's law of independent assortment refers to genes, not chromosomes, and a single chromosome may carry more than 1,000 genes. When genes are located in close proximity on the same chromosome, their alleles tend to be inherited together. This results in offspring ratios that violate Mendel's law of independent assortment. However, recombination serves to exchange genetic material on homologous chromosomes such that maternal and paternal alleles may be recombined on the same chromosome. This is why alleles on a given chromosome are not always inherited together. Recombination is a random event occurring anywhere on a chromosome. Therefore, genes that are far apart on the same chromosome are likely to still assort independently because of recombination events that occurred in the intervening chromosomal space.

Whether or not they are sorting independently, genes may interact at the level of gene products such that the expression of an allele for one gene masks or modifies the expression of an allele for a different gene. This is called epistasis.

Visual Connection Questions

[\[link\]](#) In pea plants, purple flowers (P) are dominant to white flowers (p) and yellow peas (Y) are dominant to green peas (y). What are the possible genotypes and phenotypes for a cross between PpYY and ppYy pea plants? How many squares do you need to do a Punnett square analysis of this cross?

[\[link\]](#) The possible genotypes are PpYY, PpYy, ppYY, and ppYy. The former two genotypes would result in plants with purple flowers and yellow peas, while the latter two genotypes would result in plants with white flowers with yellow peas, for a 1:1 ratio of each phenotype. You only need a 2×2 Punnett square (four squares total) to do this analysis because two of the alleles are homozygous.

Multiple Choice

Assuming no gene linkage, in a dihybrid cross of $AABB \times aabb$ with $AaBb$ F₁ heterozygotes, what is the ratio of the F₁ gametes (AB , aB , Ab , ab) that will give rise to the F₂ offspring?

1. 1:1:1:1

2. 1:3:3:1
 3. 1:2:2:1
 4. 4:3:2:1
-

A

The forked line and probability methods make use of what probability rule?

1. test cross
 2. product rule
 3. monohybrid rule
 4. sum rule
-

B

How many different offspring genotypes are expected in a trihybrid cross between parents heterozygous for all three traits when the traits behave in a dominant and recessive pattern?
How many phenotypes?

1. 64 genotypes; 16 phenotypes
 2. 16 genotypes; 64 phenotypes
 3. 8 genotypes; 27 phenotypes
 4. 27 genotypes; 8 phenotypes
-

D

Labrador retrievers' fur color is controlled by two alleles, E and B. Any dog with the ee_ genotype develops into a yellow lab, while B_E_ dogs become black labs and bbE_ dogs become chocolate labs. This is an example of ____.

1. epistasis
2. codominance
3. incomplete dominance
4. linkage

A

Which of the following situations does **not** follow the Law of Independent Assortment?

1. A blond man and a brunette woman produce three offspring over time, all of who have blond hair.
 2. A white cow crossed with a brown bull produces roan cattle.
 3. Mating a hog with a sow produces six female piglets.
 4. Men are more likely to experience hemophilia than women.
-

Critical Thinking Questions

Use the probability method to calculate the genotypes and genotypic proportions of a cross between *AABBCc* and *Aabbcc* parents.

Considering each gene separately, the cross at *A* will produce offspring of which half are *AA* and half are *Aa*; *B* will produce all *Bb*; *C* will produce half *Cc* and half *cc*. Proportions then are $(1/2) \times (1) \times (1/2)$, or $1/4$ *AABbCc*; continuing for the other possibilities yields $1/4$ *AABbcc*, $1/4$ *AaBbCc*, and $1/4$ *AaBbcc*. The proportions therefore are 1:1:1:1.

Explain epistasis in terms of its Greek-language roots “standing upon.”

Epistasis describes an antagonistic interaction between genes wherein one gene masks or interferes with the expression of another. The gene that is interfering is referred to as epistatic, as if it is “standing upon” the other

(hypostatic) gene to block its expression.

In Section 12.3, “Laws of Inheritance,” an example of epistasis was given for the summer squash. Cross white $WwYy$ heterozygotes to prove the phenotypic ratio of 12 white:3 yellow:1 green that was given in the text.

The cross can be represented as a 4×4 Punnett square, with the following gametes for each parent: WY , Wy , wY , and wy . For all 12 of the offspring that express a dominant W gene, the offspring will be white. The three offspring that are homozygous recessive for w but express a dominant Y gene will be yellow. The remaining $wwyy$ offspring will be green.

People with trisomy 21 develop Down’s syndrome. What law of Mendelian inheritance is violated in this disease? What is the most likely way this occurs?

In any trisomy disorder, a patient inherits 3 copies of a chromosome instead of the normal pair. This violates the Law of Segregation, and usually occurs when the chromosomes fail to separate during the first round of meiosis.

A heterozygous pea plant produces violet flowers and yellow, round seeds. Describe the expected genotypes of the gametes produced by Mendelian inheritance. If all three genes are found on the same arm of one chromosome should a scientist predict that inheritance patterns will follow Mendelian genetics?

Mendelian inheritance would predict that all three genes are inherited independently. There are therefore 8 different gamete genotype possibilities: VYR, VYr, VyR, Vyr, vYR, vYr, vyR, vyr. If all three genes are found on the same chromosome arm, independent assortment is unlikely to occur because the genes are close together (linked).

Glossary

dihybrid

result of a cross between two true-breeding parents that express different traits for two characteristics

epistasis

antagonistic interaction between genes such that one gene masks or interferes with the expression of another

law of dominance

in a heterozygote, one trait will conceal the presence of another trait for the same characteristic

law of independent assortment

genes do not influence each other with regard to sorting of alleles into gametes; every possible combination of alleles is equally likely to occur

law of segregation

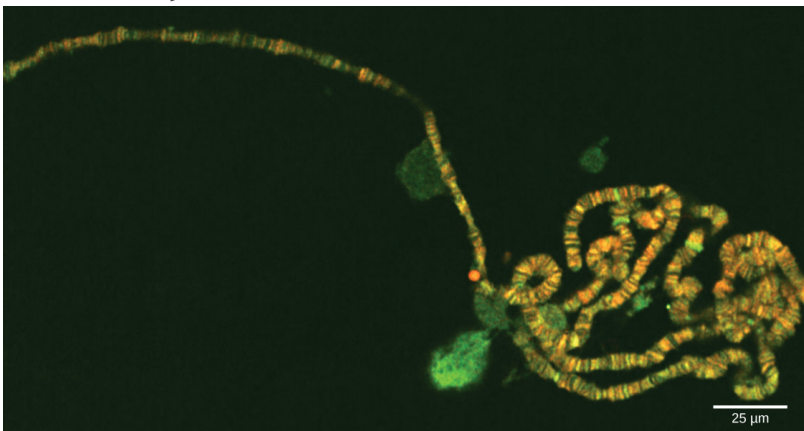
paired unit factors (i.e., genes) segregate equally into gametes such that offspring have an equal likelihood of inheriting any combination of factors

linkage

phenomenon in which alleles that are located in close proximity to each other on the same chromosome are more likely to be inherited together

Introduction

class = "introduction" Chromosomes are threadlike nuclear structures consisting of DNA and proteins that serve as the repositories for genetic information. The chromosomes depicted here were isolated from a fruit fly's salivary gland, stained with dye, and visualized under a microscope. Akin to miniature bar codes, chromosomes absorb different dyes to produce characteristic banding patterns, which allows for their routine identification. (credit: modification of work by "LPLT"/Wikimedia Commons; scale-bar data from Matt Russell)



The gene is the physical unit of inheritance, and genes are arranged in a linear order on chromosomes. Chromosome behavior and interaction during meiosis explain, at a cellular level, inheritance patterns that we observe in populations. Genetic disorders involving alterations in chromosome number or structure may have

dramatic effects and can prevent a fertilized egg from developing.

Chromosomal Theory and Genetic Linkage

By the end of this section, you will be able to do the following:

- Discuss Sutton's Chromosomal Theory of Inheritance
- Describe genetic linkage
- Explain the process of homologous recombination, or crossing over
- Describe chromosome creation
- Calculate the distances between three genes on a chromosome using a three-point test cross

Long before scientists visualized chromosomes under a microscope, the father of modern genetics, Gregor Mendel, began studying heredity in 1843. With improved microscopic techniques during the late 1800s, cell biologists could stain and visualize subcellular structures with dyes and observe their actions during cell division and meiosis. With each mitotic division, chromosomes replicated, condensed from an amorphous (no constant shape) nuclear mass into distinct X-shaped bodies (pairs of identical sister chromatids), and migrated to separate cellular poles.

(a) Walter Sutton and (b) Theodor Boveri developed the Chromosomal Theory of Inheritance, which states that chromosomes carry the unit of heredity (genes).

Chromosomal Theory of Inheritance

The speculation that chromosomes might be the key to understanding heredity led several scientists to examine Mendel's publications and reevaluate his model in terms of chromosome behavior during mitosis and meiosis. In 1902, Theodor Boveri observed that proper sea urchin embryonic development does not occur unless chromosomes are present. That same year, Walter Sutton observed chromosome separation into daughter cells during meiosis ([\[link\]](#)). Together, these observations led to the **Chromosomal Theory of Inheritance**, which identified chromosomes as the genetic material responsible for Mendelian inheritance.



(a)



(b)

The Chromosomal Theory of Inheritance was consistent with Mendel's laws, which the following observations supported:

- During meiosis, homologous chromosome pairs migrate as discrete structures that are independent of other chromosome pairs.
- Chromosome sorting from each homologous pair into pre-gametes appears to be random.
- Each parent synthesizes gametes that contain only half their chromosomal complement.
- Even though male and female gametes (sperm and egg) differ in size and morphology, they have the same number of chromosomes, suggesting equal genetic contributions from each parent.
- The gametic chromosomes combine during fertilization to produce offspring with the same chromosome number as their parents.

Despite compelling correlations between chromosome behavior during meiosis and Mendel's abstract laws, scientists proposed the Chromosomal Theory of Inheritance long before there was any direct evidence that chromosomes carried traits. Critics pointed out that individuals had far more independently segregating traits than they had chromosomes. It was only after several years of carrying out crosses with the fruit fly, *Drosophila melanogaster*, that Thomas Hunt Morgan provided experimental evidence to support the Chromosomal Theory of Inheritance.

Genetic Linkage and Distances

Mendel's work suggested that traits are inherited independently of each other. Morgan identified a 1:1 correspondence between a segregating trait and the X chromosome, suggesting that random chromosome segregation was the physical basis of Mendel's model. This also demonstrated that linked genes disrupt Mendel's predicted outcomes. That each chromosome can carry many linked genes explains how individuals can have many more traits than they have chromosomes. However, researchers in Morgan's laboratory suggested that alleles positioned on the same chromosome were not always inherited together. During meiosis, linked genes somehow became unlinked.

Homologous Recombination

In 1909, Frans Janssen observed chiasmata—the point at which chromatids are in contact with each other and may exchange segments—prior to the first meiosis division. He suggested that alleles become unlinked and chromosomes physically exchange segments. As chromosomes condensed and paired with their homologs, they appeared to interact at distinct points. Janssen suggested that these points corresponded to regions in which chromosome segments exchanged. We now know that the pairing and interaction between homologous chromosomes, or synapsis, does more than simply organize the homologs for migration to separate daughter cells. When synapsed, homologous chromosomes undergo

reciprocal physical exchanges at their arms in **homologous recombination**, or more simply, “crossing over.”

To better understand the type of experimental results that researchers were obtaining at this time, consider a heterozygous individual that inherited dominant maternal alleles for two genes on the same chromosome (such as AB) and two recessive paternal alleles for those same genes (such as ab). If the genes are linked, one would expect this individual to produce gametes that are either AB or ab with a 1:1 ratio. If the genes are unlinked, the individual should produce AB , Ab , aB , and ab gametes with equal frequencies, according to the Mendelian concept of independent assortment. Because they correspond to new allele combinations, the genotypes Ab and aB are **nonparental types** that result from homologous recombination during meiosis. **Parental types** are progeny that exhibit the same allelic combination as their parents. Morgan and his colleagues, however, found that when they test crossed such heterozygous individuals to a homozygous recessive parent ($AaBb \times aabb$), both parental and nonparental cases occurred. For example, 950 offspring might be recovered that were either $AaBb$ or $aabb$, but 50 offspring would also result that were either $Aabb$ or $aaBb$. These results suggested that linkage occurred most often, but a significant minority of offspring were the products of recombination.

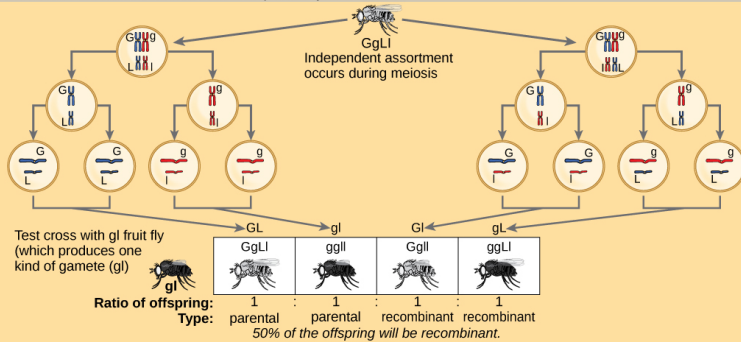
Visual Connection

This figure shows unlinked and linked gene inheritance patterns. In (a), two genes are located on different chromosomes so independent assortment occurs during meiosis. The offspring have an equal chance of being the parental type (inheriting the same combination of traits as the parents) or a nonparental type (inheriting a different combination of traits than the parents). In (b), two genes are very close together on the same chromosome so that no crossing over occurs between them. Therefore, the genes are always inherited together and all the offspring are the parental type. In (c), two genes are far apart on the chromosome such that crossing over occurs during every meiotic event. The recombination frequency will be the same as if the genes were on separate chromosomes. (d) The actual recombination frequency of fruit fly wing length and body color that Thomas Morgan observed in 1912 was 17 percent. A crossover frequency between 0 percent and 50 percent indicates that the genes are on the same chromosome and crossover sometimes occurs.

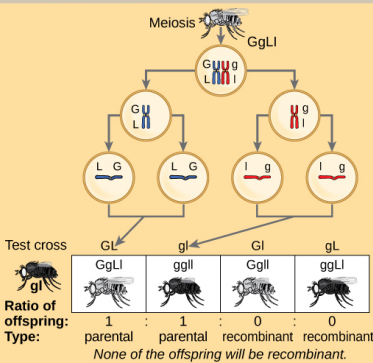
Inheritance Pattern of Linked and Unlinked Genes

Three hypothetical inheritance patterns for a test cross between a heterozygote and a homozygous recessive individual, based on gene placement, are shown in A through C. The actual experimental results published by Thomas Hunt Morgan in 1912 are shown in D.

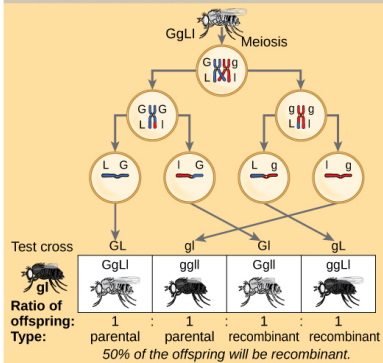
A. Genes on different chromosomes, independently assorted



B. Genes on the same chromosome, no crossover occurs



C. Genes on the same chromosome, crossover occurs 100% of the time



D. Results from Morgan's 1912 experiment

Test cross	GL	gl	Gl	gL
	Gg Ll	gg ll	Gg ll	gg Ll
Number of offspring:	965	944	206	185
Ratio of offspring:	1	1	.2	.2
Type:	parental	parental	recombinant	recombinant

17% of the offspring are recombinant, indicating that the genes are on the same chromosome and crossover occurs some of the time.

In a test cross for two characteristics such as the one here, can the recombinant offspring's predicted frequency be 60 percent? Why or why not?

Genetic Maps

Janssen did not have the technology to demonstrate

crossing over so it remained an abstract idea that scientists did not widely believe. Scientists thought chiasmata were a variation on synapsis and could not understand how chromosomes could break and rejoin. Yet, the data were clear that linkage did not always occur. Ultimately, it took a young undergraduate student and an “all-nighter” to mathematically elucidate the linkage and recombination problem.

In 1913, Alfred Sturtevant, a student in Morgan’s laboratory, gathered results from researchers in the laboratory, and took them home one night to mull them over. By the next morning, he had created the first “chromosome map,” a linear representation of gene order and relative distance on a chromosome ([\[link\]](#)).

Visual Connection

This genetic map orders *Drosophila* genes on the basis of recombination frequency.

Genetic Map Based on Recombination Frequencies in <i>Drosophila</i>				
MUTANT				WILD TYPE
Short aristae	0			Long aristae
Black body	48.5			Gray body
Cinnabar eyes	57.5			Red eyes
Vestigial wings	65.5			Normal wings
Brown eyes	104.5			Red eyes
Values in centimorgan (cM) map units; recombination frequency of 0.01 = 1 cM				

Which of the following statements is true?

1. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.
2. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
3. Recombination of the gray/black body color and long/short aristae alleles will not occur.
4. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.

As [\[link\]](#) shows, by using recombination frequency to predict genetic distance, we can infer the relative gene order on chromosome 2. The values represent map distances in centimorgans (cM), which correspond to recombination frequencies (in percent). Therefore, the genes for body color and wing size were $65.5 - 48.5 = 17$ cM apart, indicating that the maternal and paternal alleles for these genes recombine in 17 percent of offspring, on average.

To construct a chromosome map, Sturtevant assumed that genes were ordered serially on threadlike chromosomes. He also assumed that the incidence of recombination between two homologous chromosomes could occur with equal likelihood anywhere along the chromosome's length. Operating under these assumptions, Sturtevant postulated that alleles that were far apart on a chromosome were more likely to dissociate during meiosis simply because there was a larger region over which recombination could occur. Conversely, alleles that were close to each other on the chromosome were likely to be inherited together. The average number of crossovers between two alleles—that is, their **recombination frequency**—correlated with their genetic distance from each other, relative to the locations of other genes on that chromosome. Considering the example cross between *AaBb* and *aabb* above, we could calculate the recombination's frequency as $50/1000 = 0.05$.

That is, the likelihood of a crossover between genes A/a and B/b was 0.05, or 5 percent. Such a result would indicate that the genes were definitively linked, but that they were far enough apart for crossovers to occasionally occur. Sturtevant divided his genetic map into map units, or **centimorgans (cM)**, in which a 0.01 recombination frequency corresponds to 1 cM.

By representing alleles in a linear map, Sturtevant suggested that genes can range from linking perfectly (recombination frequency = 0) to unlinking perfectly (recombination frequency = 0.5) when genes are on different chromosomes or genes separate very far apart on the same chromosome. Perfectly unlinked genes correspond to the frequencies Mendel predicted to assort independently in a dihybrid cross. A 0.5 recombination frequency indicates that 50 percent of offspring are recombinants and the other 50 percent are parental types. That is, every type of allele combination is represented with equal frequency. This representation allowed Sturtevant to additively calculate distances between several genes on the same chromosome. However, as the genetic distances approached 0.50, his predictions became less accurate because it was not clear whether the genes were very far apart on the same or on different chromosomes.

In 1931, Barbara McClintock and Harriet Creighton

demonstrated the crossover of homologous chromosomes in corn plants. Weeks later, Curt Stern demonstrated microscopically homologous recombination in *Drosophila*. Stern observed several X-linked phenotypes that were associated with a structurally unusual and dissimilar X chromosome pair in which one X was missing a small terminal segment, and the other X was fused to a piece of the Y chromosome. By crossing flies, observing their offspring, and then visualizing the offspring's chromosomes, Stern demonstrated that every time the offspring allele combination deviated from either of the parental combinations, there was a corresponding exchange of an X chromosome segment. Using mutant flies with structurally distinct X chromosomes was the key to observing the products of recombination because DNA sequencing and other molecular tools were not yet available. We now know that homologous chromosomes regularly exchange segments in meiosis by reciprocally breaking and rejoining their DNA at precise locations.

Link to Learning

Review Sturtevant's process to create a genetic map on the basis of recombination frequencies [here](#).

Mendel's Mapped Traits

Homologous recombination is a common genetic process, yet Mendel never observed it. Had he investigated both linked and unlinked genes, it would have been much more difficult for him to create a unified model of his data on the basis of probabilistic calculations. Researchers who have since mapped the seven traits that Mendel investigated onto a pea plant genome's seven chromosomes have confirmed that all the genes he examined are either on separate chromosomes or are sufficiently far apart as to be statistically unlinked. Some have suggested that Mendel was enormously lucky to select only unlinked genes; whereas, others question whether Mendel discarded any data suggesting linkage. In any case, Mendel consistently observed independent assortment because he examined genes that were effectively unlinked.

Section Summary

Sutton and Boveri's Chromosomal Theory of Inheritance states that chromosomes are the vehicles of genetic heredity. Neither Mendelian genetics nor gene linkage is perfectly accurate. Instead, chromosome behavior involves segregation, independent assortment, and occasionally, linkage. Sturtevant devised a method to assess

recombination frequency and infer linked genes' relative positions and distances on a chromosome on the basis of the average number of crossovers in the intervening region between the genes. Sturtevant correctly presumed that genes are arranged in serial order on chromosomes and that recombination between homologs can occur anywhere on a chromosome with equal likelihood. Whereas linkage causes alleles on the same chromosome to be inherited together, homologous recombination biases alleles toward an independent inheritance pattern.

Visual Connection Questions

[\[link\]](#) In a test cross for two characteristics such as the one shown here, can the predicted frequency of recombinant offspring be 60 percent? Why or why not?

[\[link\]](#) No. The predicted frequency of recombinant offspring ranges from 0% (for linked traits) to 50% (for unlinked traits).

[\[link\]](#) Which of the following statements is true?

1. Recombination of the body color and red/cinnabar eye alleles will occur more frequently than recombination of the alleles for wing length and aristae length.
2. Recombination of the body color and aristae length alleles will occur more frequently than recombination of red/brown eye alleles and the aristae length alleles.
3. Recombination of the gray/black body color and long/short aristae alleles will not occur.
4. Recombination of the red/brown eye and long/short aristae alleles will occur more frequently than recombination of the alleles for wing length and body color.

[\[link\]](#) D

Review Questions

X-linked recessive traits in humans (or in *Drosophila*) are observed _____.

1. in more males than females
2. in more females than males
3. in males and females equally

4. in different distributions depending on the trait

A

The first suggestion that chromosomes may physically exchange segments came from the microscopic identification of _____.

1. synapsis
 2. sister chromatids
 3. chiasmata
 4. alleles
-

C

Which recombination frequency corresponds to independent assortment and the absence of linkage?

1. 0
 2. 0.25
 3. 0.50
 4. 0.75
-

C

Which recombination frequency corresponds to perfect linkage and violates the law of independent assortment?

1. 0
2. 0.25
3. 0.50
4. 0.75

A

Critical Thinking Questions

Explain how the Chromosomal Theory of Inheritance helped to advance our understanding of genetics.

The Chromosomal Theory of Inheritance proposed that genes reside on chromosomes. The understanding that chromosomes are linear arrays of genes explained linkage, and crossing over explained recombination.

Glossary

centimorgan (cM)

(also, map unit) relative distance that corresponds to a 0,01 recombination frequency

Chromosomal Theory of Inheritance

theory proposing that chromosomes are the genes' vehicles and that their behavior during meiosis is the physical basis of the inheritance patterns that Mendel observed

homologous recombination

process by which homologous chromosomes undergo reciprocal physical exchanges at their arms, also crossing over

nonparental (recombinant) type

progeny resulting from homologous recombination that exhibits a different allele combination compared with its parents

parental types

progeny that exhibits the same allelic combination as its parents

recombination frequency

average number of crossovers between two alleles; observed as the number of nonparental types in a progeny's population

Chromosomal Basis of Inherited Disorders

By the end of this section, you will be able to do the following:

- Describe how a karyogram is created
- Explain how nondisjunction leads to disorders in chromosome number
- Compare disorders that aneuploidy causes
- Describe how errors in chromosome structure occur through inversions and translocations

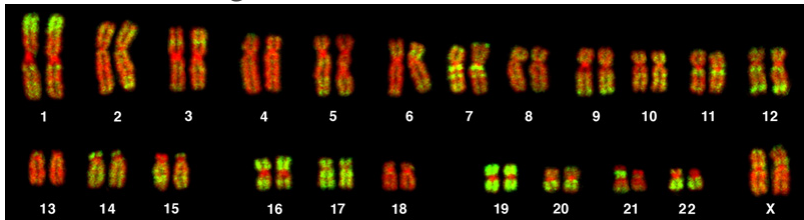
Inherited disorders can arise when chromosomes behave abnormally during meiosis. We can divide chromosome disorders into two categories: abnormalities in chromosome number and chromosomal structural rearrangements. Because even small chromosome segments can span many genes, chromosomal disorders are characteristically dramatic and often fatal.

This karyotype is of a female human. Notice that homologous chromosomes are the same size, and have the same centromere positions and banding patterns. A human male would have an XY chromosome pair instead of the XX pair. (credit: Andreas Blozer et al)

Chromosome Identification

Chromosome isolation and microscopic observation forms the basis of cytogenetics and is the primary

method by which clinicians detect chromosomal abnormalities in humans. A **karyotype** is the number and appearance of chromosomes, and includes their length, banding pattern, and centromere position. To obtain a view of an individual's karyotype, cytologists photograph the chromosomes and then cut and paste each chromosome into a chart, or **karyogram**. Another name is an ideogram ([link](#)).



In a given species, we can identify chromosomes by their number, size, centromere position, and banding pattern. In a human karyotype, **autosomes** or “body chromosomes” (all of the non-sex chromosomes) are generally organized in approximate order of size from largest (chromosome 1) to smallest (chromosome 22). The X and Y chromosomes are not autosomes. However, chromosome 21 is actually shorter than chromosome 22. Researchers discovered this after naming Down syndrome as trisomy 21, reflecting how this disease results from possessing one extra chromosome 21 (three total). Not wanting to change the name of this important disease, scientists retained the numbering of chromosome 21 despite describing it having the shortest set of

chromosomes. We may designate the chromosome “arms” projecting from either end of the centromere as short or long, depending on their relative lengths. We abbreviate the short arm p (for “petite”); whereas, we abbreviate the long arm q (because it follows “p” alphabetically). Numbers further subdivide and denote each arm. Using this naming system, we can describe chromosome locations consistently in the scientific literature.

Career Connection

Geneticists Use Karyograms to Identify Chromosomal Aberrations

Although we refer to Mendel as the “father of modern genetics,” he performed his experiments with none of the tools that the geneticists of today routinely employ. One such powerful cytological technique is karyotyping, a method in which geneticists can identify traits characterized by chromosomal abnormalities from a single cell. To observe an individual’s karyotype, a geneticist first collects a person’s cells (like white blood cells) from a blood sample or other tissue. In the laboratory, he or she stimulates the isolated cells to begin actively dividing. The geneticist then applies the chemical colchicine to cells to arrest condensed chromosomes in metaphase. The geneticist then induces swelling in the cells using a hypotonic solution so the chromosomes spread apart. Finally,

the geneticist preserves the sample in a fixative and applies it to a slide.

The geneticist then stains chromosomes with one of several dyes to better visualize each chromosome pair's distinct and reproducible banding patterns. Following staining, the geneticist views the chromosomes using bright-field microscopy. A common stain choice is the Giemsa stain. Giemsa staining results in approximately 400–800 bands (of tightly coiled DNA and condensed proteins) arranged along all 23 chromosome pairs. An experienced geneticist can identify each band. In addition to the banding patterns, geneticists further identify chromosomes on the basis of size and centromere location. To obtain the classic depiction of the karyotype in which homologous chromosome pairs align in numerical order from longest to shortest, the geneticist obtains a digital image, identifies each chromosome, and manually arranges the chromosomes into this pattern ([\[link\]](#)).

At its most basic, the karyogram may reveal genetic abnormalities in which an individual has too many or too few chromosomes per cell. Examples of this are Down Syndrome, which one identifies by a third copy of chromosome 21, and Turner Syndrome, which has the presence of only one X chromosome in women instead of the normal two characterizes. Geneticists can also identify large DNA deletions or insertions. For instance, geneticists can identify Jacobsen Syndrome—which

involves distinctive facial features as well as heart and bleeding defects—by a deletion on chromosome 11. Finally, the karyotype can pinpoint **translocations**, which occur when a segment of genetic material breaks from one chromosome and reattaches to another chromosome or to a different part of the same chromosome. Translocations are implicated in certain cancers, including chronic myelogenous leukemia.

During Mendel's lifetime, inheritance was an abstract concept that one could only infer by performing crosses and observing the traits that offspring expressed. By observing a karyogram, today's geneticists can actually visualize an individual's chromosomal composition to confirm or predict genetic abnormalities in offspring, even before birth.

Chromosome Number Disorders

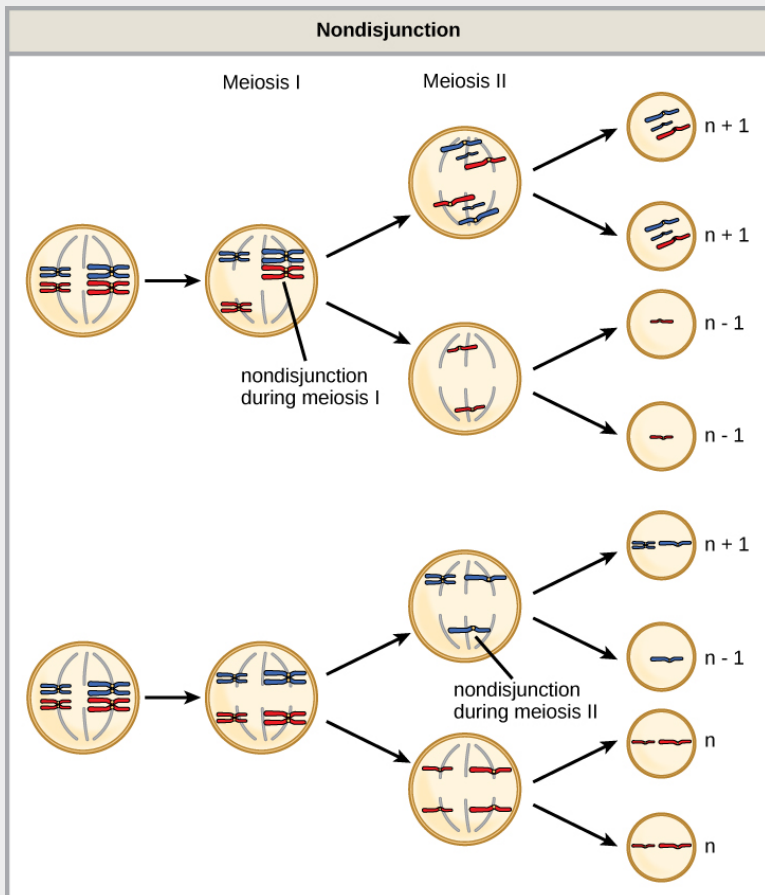
Of all of the chromosomal disorders, chromosome number abnormalities are the most obviously identifiable from a karyogram. Chromosome number disorders include duplicating or losing entire chromosomes, as well as changes in the number of complete sets of chromosomes. They are

caused by **nondisjunction**, which occurs when homologous chromosome pairs or sister chromatids fail to separate during meiosis. Misaligned or incomplete synapsis, or a spindle apparatus dysfunction that facilitates chromosome migration, can cause nondisjunction. The risk of nondisjunction occurring increases with the parents' age.

Nondisjunction can occur during either meiosis I or II, with differing results ([\[link\]](#)). If homologous chromosomes fail to separate during meiosis I, the result is two gametes that lack that particular chromosome and two gametes with two chromosome copies. If sister chromatids fail to separate during meiosis II, the result is one gamete that lacks that chromosome, two normal gametes with one chromosome copy, and one gamete with two chromosome copies.

Visual Connection

Nondisjunction occurs when homologous chromosomes or sister chromatids fail to separate during meiosis, resulting in an abnormal chromosome number. Nondisjunction may occur during meiosis I or meiosis II.



Which of the following statements about nondisjunction is true?

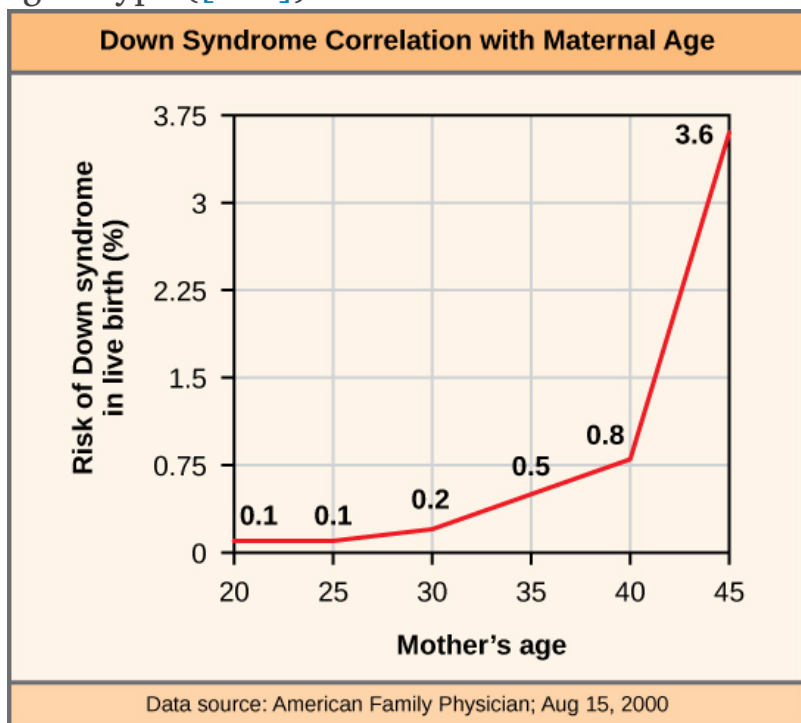
1. Nondisjunction only results in gametes with $n+1$ or $n-1$ chromosomes.
2. Nondisjunction occurring during meiosis II results in 50 percent normal gametes.
3. Nondisjunction during meiosis I results in 50 percent normal gametes.
4. Nondisjunction always results in four different kinds of gametes.

The incidence of having a fetus with trisomy 21 increases dramatically with maternal age.

Aneuploidy

Scientists call an individual with the appropriate number of chromosomes for their species **euploid**. In humans, euploidy corresponds to 22 pairs of autosomes and one pair of sex chromosomes. An individual with an error in chromosome number is described as **aneuploid**, a term that includes **monosomy** (losing one chromosome) or **trisomy** (gaining an extraneous chromosome). Monosomic human zygotes missing any one copy of an autosome invariably fail to develop to birth because they lack essential genes. This underscores the importance of “gene dosage” in humans. Most autosomal trisomies also fail to develop to birth; however, duplications of some smaller chromosomes (13, 15, 18, 21, or 22) can result in offspring that survive for several weeks to many years. Trisomic individuals suffer from a different type of genetic imbalance: an excess in gene dose. Individuals with an extra chromosome may synthesize an abundance of the gene products, which that chromosome encodes. This extra dose (150 percent) of specific genes can lead to a number of functional challenges and often precludes development. The most common trisomy among viable births is that of chromosome 21, which corresponds to Down Syndrome. Short stature and stunted digits, facial

distinctions that include a broad skull and large tongue, and significant developmental delays characterize individuals with this inherited disorder. We can correlate the incidence of Down syndrome with maternal age. Older women are more likely to become pregnant with fetuses carrying the trisomy 21 genotype ([\[link\]](#)).



Link to Learning

Visualize adding a chromosome that leads to Down syndrome in this [video simulation](#).

As with many polyploid plants, this triploid orange daylily (*Hemerocallis fulva*) is particularly large and robust, and grows flowers with triple the number of petals of its diploid counterparts. (credit: Steve Karg)

Polyploidy

We call an individual with more than the correct number of chromosome sets (two for diploid species) **polyploid**. For instance, fertilizing an abnormal diploid egg with a normal haploid sperm would yield a triploid zygote. Polyploid animals are extremely rare, with only a few examples among the flatworms, crustaceans, amphibians, fish, and lizards. Polyploid animals are sterile because meiosis cannot proceed normally and instead produces mostly aneuploid daughter cells that cannot yield viable zygotes. Rarely, polyploid animals can reproduce asexually by haplodiploidy, in which an unfertilized egg divides mitotically to produce offspring. In contrast, polyploidy is very common in the plant kingdom, and polyploid plants tend to be larger and more robust than euploids of their species ([\[link\]](#)).



In cats, the gene for coat color is located on the X chromosome. In female cats' embryonic development, one of the two X chromosomes randomly inactivates in each cell, resulting in a tortoiseshell pattern if the cat has two different alleles for coat color. Male cats, having only one X chromosome, never exhibit a tortoiseshell coat color. (credit: Michael Bodega)

Sex Chromosome Nondisjunction in Humans

Humans display dramatic deleterious effects with autosomal trisomies and monosomies. Therefore, it may seem counterintuitive that human females and males can function normally, despite carrying different numbers of the X chromosome. Rather

than a gain or loss of autosomes, variations in the number of sex chromosomes occur with relatively mild effects. In part, this happens because of the molecular process **X inactivation**. Early in development, when female mammalian embryos consist of just a few thousand cells (relative to trillions in the newborn), one X chromosome in each cell inactivates by tightly condensing into a quiescent (dormant) structure, or a Barr body. The chance that an X chromosome (maternally or paternally derived) inactivates in each cell is random, but once this occurs, all cells derived from that one will have the same inactive X chromosome or Barr body. By this process, females compensate for their double genetic dose of X chromosome. In so-called “tortoiseshell” cats, we observe embryonic X inactivation as color variegation ([\[link\]](#)). Females that are heterozygous for an X-linked coat color gene will express one of two different coat colors over different regions of their body, corresponding to whichever X chromosome inactivates in that region's embryonic cell progenitor.



An individual carrying an abnormal number of X chromosomes will inactivate all but one X chromosome in each of her cells. However, even inactivated X chromosomes continue to express a few genes, and X chromosomes must reactivate for the proper maturation of female ovaries. As a result, X-chromosomal abnormalities typically occur with mild mental and physical defects, as well as sterility. If the X chromosome is absent altogether, the individual will not develop in utero.

Scientists have identified and characterized several errors in sex chromosome number. Individuals with three X chromosomes, triplo-X, are phenotypically female but express developmental delays and reduced fertility. The XXY genotype, corresponding

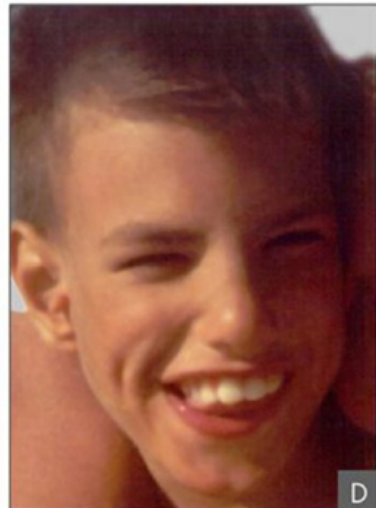
to one type of Klinefelter syndrome, corresponds to phenotypically male individuals with small testes, enlarged breasts, and reduced body hair. More complex types of Klinefelter syndrome exist in which the individual has as many as five X chromosomes. In all types, every X chromosome except one undergoes inactivation to compensate for the excess genetic dosage. We see this as several Barr bodies in each cell nucleus. Turner syndrome, characterized as an XO genotype (i.e., only a single sex chromosome), corresponds to a phenotypically female individual with short stature, webbed skin in the neck region, hearing and cardiac impairments, and sterility.

This figure shows an individual with cri-du-chat syndrome at two, four, nine, and 12 years of age. (credit: Paola Cerruti Mainardi)

Duplications and Deletions

In addition to losing or gaining an entire chromosome, a chromosomal segment may duplicate or lose itself. Duplications and deletions often produce offspring that survive but exhibit physical and mental abnormalities. Duplicated chromosomal segments may fuse to existing chromosomes or may be free in the nucleus. Cri-du-chat (from the French for “cry of the cat”) is a syndrome that occurs with nervous system abnormalities and identifiable physical features that result from a deletion of most 5p (the small arm of

chromosome 5) ([\[link\]](#)). Infants with this genotype emit a characteristic high-pitched cry on which the disorder's name is based.



Pericentric inversions include the centromere, and paracentric inversions do not. A pericentric inversion can change the chromosome arms' relative lengths. A paracentric inversion cannot. When one chromosome undergoes an inversion but the other

does not, one chromosome must form an inverted loop to retain point-for-point interaction during synapsis. This inversion pairing is essential to maintaining gene alignment during meiosis and to allow for recombination. A reciprocal translocation occurs when a DNA segment transfers from one chromosome to another, nonhomologous chromosome. (credit: modification of work by National Human Genome Research/USA)

Chromosomal Structural Rearrangements

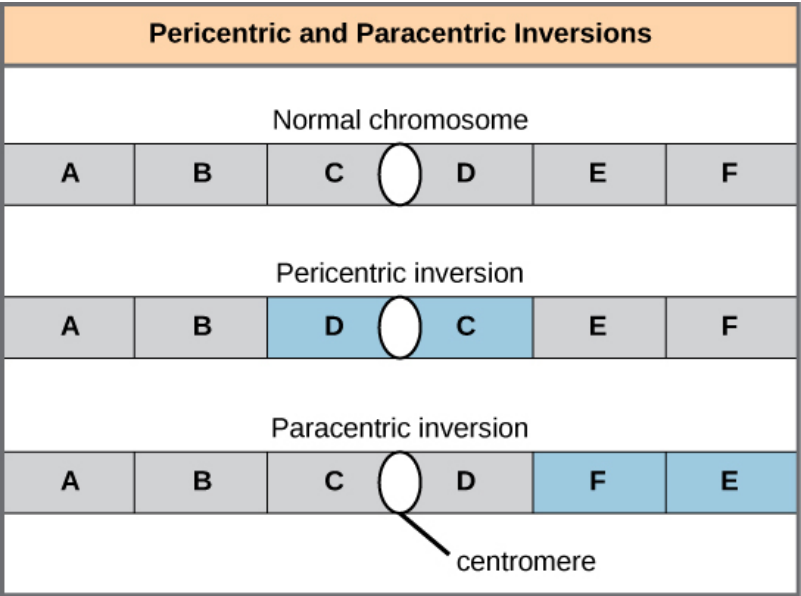
Cytologists have characterized numerous structural rearrangements in chromosomes, but chromosome inversions and translocations are the most common. We can identify both during meiosis by the adaptive pairing of rearranged chromosomes with their former homologs to maintain appropriate gene alignment. If the genes on two homologs are not oriented correctly, a recombination event could result in losing genes from one chromosome and gaining genes on the other. This would produce aneuploid gametes.

Chromosome Inversions

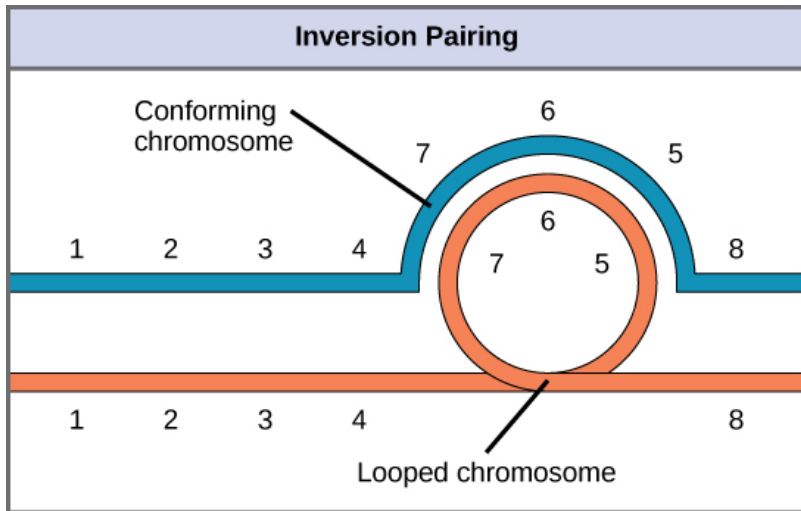
A **chromosome inversion** is the detachment, 180° rotation, and reinsertion of part of a chromosome. Inversions may occur in nature as a result of mechanical shear, or from transposable elements'

action (special DNA sequences capable of facilitating rearranging chromosome segments with the help of enzymes that cut and paste DNA sequences). Unless they disrupt a gene sequence, inversions only change gene orientation and are likely to have more mild effects than aneuploid errors. However, altered gene orientation can result in functional changes because regulators of gene expression could move out of position with respect to their targets, causing aberrant levels of gene products.

An inversion can be **pericentric** and include the centromere, or **paracentric** and occur outside the centromere ([link](#)). A pericentric inversion that is asymmetric about the centromere can change the chromosome arms' relative lengths, making these inversions easily identifiable.



When one homologous chromosome undergoes an inversion but the other does not, the individual is an inversion heterozygote. To maintain point-for-point synapsis during meiosis, one homolog must form a loop, and the other homolog must mold around it. Although this topology can ensure that the genes correctly align, it also forces the homologs to stretch and can occur with imprecise synapsis regions ([\[link\]](#)).



Evolution Connection

The Chromosome 18 Inversion

Not all chromosomes' structural rearrangements produce nonviable, impaired, or infertile individuals. In rare instances, such a change can result in new species evolving. In fact, a pericentric inversion in chromosome 18 appears to have contributed to human evolution. This inversion is

not present in our closest genetic relatives, the chimpanzees. Humans and chimpanzees differ cytogenetically by pericentric inversions on several chromosomes and by the fusion of two separate chromosomes in chimpanzees that correspond to chromosome two in humans.

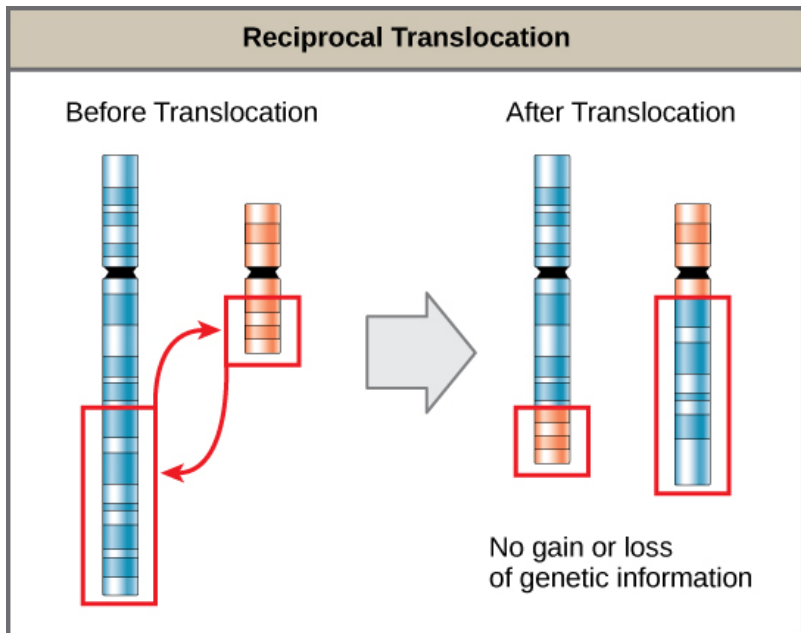
Scientists believe the pericentric chromosome 18 inversion occurred in early humans following their divergence from a common ancestor with chimpanzees approximately five million years ago. Researchers characterizing this inversion have suggested that approximately 19,000 nucleotide bases were duplicated on 18p, and the duplicated region inverted and reinserted on chromosome 18 of an ancestral human.

A comparison of human and chimpanzee genes in the region of this inversion indicates that two genes—*ROCK1* and *USP14*—that are adjacent on chimpanzee chromosome 17 (which corresponds to human chromosome 18) are more distantly positioned on human chromosome 18. This suggests that one of the inversion breakpoints occurred between these two genes. Interestingly, humans and chimpanzees express *USP14* at distinct levels in specific cell types, including cortical cells and fibroblasts. Perhaps the chromosome 18 inversion in an ancestral human repositioned specific genes and reset their expression levels in a useful way. Because both *ROCK1* and *USP14* encode cellular enzymes, a change in their expression could alter cellular function. We do not

know how this inversion contributed to hominid evolution, but it appears to be a significant factor in the divergence of humans from other primates. [Violaine Goidts et al.](#), “Segmental duplication associated with the human-specific inversion of chromosome 18: a further example of the impact of segmental duplications on karyotype and genome evolution in primates,” *Human Genetics*. 115 (2004):116-122

Translocations

A **translocation** occurs when a chromosome segment dissociates and reattaches to a different, nonhomologous chromosome. Translocations can be benign or have devastating effects depending on how the positions of genes are altered with respect to regulatory sequences. Notably, specific translocations have occurred with several cancers and with schizophrenia. Reciprocal translocations result from exchanging chromosome segments between two nonhomologous chromosomes such that there is no genetic information gain or loss ([\[link\]](#)).



Section Summary

The number, size, shape, and banding pattern of chromosomes make them easily identifiable in a karyogram and allows for the assessment of many chromosomal abnormalities. Disorders in chromosome number, or aneuploidies, are typically lethal to the embryo, although a few trisomic genotypes are viable. Because of X inactivation, aberrations in sex chromosomes typically have milder phenotypic effects. Aneuploidies also include instances in which a chromosome's segments duplicate or delete themselves. Inversion or translocation also may rearrange chromosome structures. Both of these aberrations can result in

problematic phenotypic effects. Because they force chromosomes to assume unnatural topologies during meiosis, inversions and translocations often occur with reduced fertility because of the likelihood of nondisjunction.

Visual Connection Questions

[\[link\]](#) Which of the following statements about nondisjunction is true?

1. Nondisjunction only results in gametes with $n + 1$ or $n - 1$ chromosomes.
2. Nondisjunction occurring during meiosis II results in 50 percent normal gametes.
3. Nondisjunction during meiosis I results in 50 percent normal gametes.
4. Nondisjunction always results in four different kinds of gametes.

[\[link\]](#) B.

Review Questions

Which of the following codes describes position 12 on the long arm of chromosome 13?

1. 13p12
2. 13q12
3. 12p13
4. 12q13

B

In agriculture, polyploid crops (like coffee, strawberries, or bananas) tend to produce _____.

1. more uniformity
2. more variety
3. larger yields
4. smaller yields

C

Assume a pericentric inversion occurred in one of two homologs prior to meiosis. The other homolog remains normal. During meiosis, what structure—if any—would these homologs assume in order to pair accurately along their lengths?

1. V formation
 2. cruciform
 3. loop
 4. pairing would not be possible
-

C

The genotype XXY corresponds to

1. Klinefelter syndrome
 2. Turner syndrome
 3. Triplo-X
 4. Jacob syndrome
-

A

Abnormalities in the number of X chromosomes tends to have milder phenotypic effects than the same abnormalities in autosomes because of _____.

1. deletions
 2. nonhomologous recombination
 3. synapsis
 4. X inactivation
-

D

By definition, a pericentric inversion includes the _____.

1. centromere
2. chiasma
3. telomere
4. synapse

A

Critical Thinking Questions

Using diagrams, illustrate how nondisjunction can result in an aneuploid zygote.

Exact diagram style will vary; diagram should look like [\[link\]](#).

Glossary

aneuploid

individual with an error in chromosome number; includes chromosome segment deletions and duplications

autosome

any of the non-sex chromosomes

chromosome inversion

detachment, 180° rotation, and chromosome arm reinsertion

euploid

individual with the appropriate number of chromosomes for their species

karyogram

a karyotype's photographic image

karyotype

an individual's chromosome number and appearance; includes the size, banding patterns, and centromere position

monosomy

otherwise diploid genotype in which one chromosome is missing

nondisjunction

failure of synapsed homologs to completely separate and migrate to separate poles during the meiosis' first cell division

paracentric

inversion that occurs outside the centromere

pericentric

inversion that involves the centromere

polyploid

individual with an incorrect number of chromosome sets

translocation

process by which one chromosome segment dissociates and reattaches to a different, nonhomologous chromosome

trisomy

otherwise diploid genotype in which one entire chromosome duplicates

X inactivation

condensing X chromosomes into Barr bodies during embryonic development in females to compensate for the double genetic dose

Introduction

class = "introduction" Dolly the sheep was the first large mammal to be cloned.



The three letters “DNA” have now become synonymous with crime solving and genetic testing. DNA can be retrieved from hair, blood, or saliva. Each person’s DNA is unique, and it is possible to detect differences between individuals within a species on the basis of these unique features.

DNA analysis has many practical applications beyond forensics. In humans, DNA testing is applied to numerous uses: determining paternity, tracing genealogy, identifying pathogens, archeological research, tracing disease outbreaks, and studying human migration patterns. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. It is now possible to determine predisposition to diseases by looking at genes.

Each human cell has 23 pairs of chromosomes: one set of chromosomes is inherited from the mother and the other set is inherited from the father. There is also a mitochondrial genome, inherited exclusively from the mother, which can be involved in inherited genetic disorders. On each chromosome, there are thousands of genes that are responsible for determining the genotype and phenotype of the individual. A gene is defined as a sequence of DNA that codes for a functional product. The human haploid genome contains 3 billion base pairs and has between 20,000 and 25,000 functional genes.

Historical Basis of Modern Understanding

By the end of this section, you will be able to do the following:

- Explain transformation of DNA
- Describe the key experiments that helped identify that DNA is the genetic material
- State and explain Chargaff's rules

Our current understanding of DNA began with the discovery of nucleic acids followed by the development of the double-helix model. In the 1860s, Friedrich Miescher ([\[link\]](#)), a physician by profession, isolated phosphate-rich chemicals from white blood cells (leukocytes). He named these chemicals (which would eventually be known as DNA) *nuclein* because they were isolated from the nuclei of the cells.

Friedrich Miescher (1844–1895) discovered nucleic acids.



Link to Learning

To see Miescher conduct his experiment that led to his discovery of DNA and associated proteins in the nucleus, click through [this review](#).

A half century later, in 1928, British bacteriologist Frederick Griffith reported the first demonstration of bacterial **transformation**—a process in which external DNA is taken up by a cell, thereby changing its morphology and physiology. Griffith conducted his experiments with *Streptococcus pneumoniae*, a bacterium that causes pneumonia. Griffith worked with two strains of this bacterium called rough (R) and smooth (S). (The two cell types were called

“rough” and “smooth” after the appearance of their colonies grown on a nutrient agar plate.)

The R strain is non-pathogenic (does not cause disease). The S strain is pathogenic (disease-causing), and has a capsule outside its cell wall. The capsule allows the cell to escape the immune responses of the host mouse.

When Griffith injected the living S strain into mice, they died from pneumonia. In contrast, when Griffith injected the live R strain into mice, they survived. In another experiment, when he injected mice with the heat-killed S strain, they also survived. This experiment showed that the capsule alone was not the cause of death. In a third set of experiments, a mixture of live R strain and heat-killed S strain were injected into mice, and—to his surprise—the mice died. Upon isolating the live bacteria from the dead mouse, only the S strain of bacteria was recovered. When this isolated S strain was injected into fresh mice, the mice died. Griffith concluded that something had passed from the heat-killed S strain into the live R strain and transformed it into the pathogenic S strain. He called this the *transforming principle* ([\[link\]](#)). These experiments are now known as Griffith's transformation experiments. Two strains of *S. pneumoniae* were used in Griffith's transformation experiments. The R strain is non-pathogenic, whereas the S strain is pathogenic and causes death. When Griffith injected a mouse with

the heat-killed S strain and a live R strain, the mouse died. The S strain was recovered from the dead mouse. Griffith concluded that something had passed from the heat-killed S strain to the R strain, transforming the R strain into the S strain in the process. (credit "living mouse": modification of work by NIH; credit "dead mouse": modification of work by Sarah Marriage)



Mouse injected with heat-killed virulent S strain lives.



Mouse injected with both heat-killed S strain and live non-virulent R strain dies.

Scientists Oswald Avery, Colin MacLeod, and Maclyn McCarty (1944) were interested in exploring this transforming principle further. They isolated the S strain from the dead mice and isolated the proteins and nucleic acids (RNA and DNA) as these were possible candidates for the molecule of heredity. They used enzymes that specifically degraded each component and then used each mixture separately to transform the R strain. They found that when DNA was degraded, the resulting mixture was no longer able to transform the bacteria, whereas all of the other combinations were able to transform the bacteria. This led them to conclude that DNA was the transforming principle.

Career Connection

Forensic Scientist

Forensic Scientists used DNA analysis evidence for the first time to solve an immigration case. The story started with a teenage boy returning to London from Ghana to be with his mother.

Immigration authorities at the airport were suspicious of him, thinking that he was traveling on a forged passport. After much persuasion, he was allowed to go live with his mother, but the immigration authorities did not drop the case against him. All types of evidence, including photographs, were provided to the authorities, but deportation proceedings were started nevertheless. Around the same time, Dr. Alec Jeffreys of Leicester University in the United Kingdom had invented a technique known as **DNA**

fingerprinting. The immigration authorities approached Dr. Jeffreys for help. He took DNA samples from the mother and three of her children, as well as an unrelated mother, and compared the samples with the boy's DNA. Because the biological father was not in the picture, DNA from the three children was compared with the boy's DNA. He found a match in the boy's DNA for both the mother and his three siblings. He concluded that the boy was indeed the mother's son.

Forensic scientists analyze many items, including documents, handwriting, firearms, and biological samples. They analyze the DNA content of hair, semen, saliva, and blood, and compare it with a

database of DNA profiles of known criminals. Analysis includes DNA isolation, sequencing, and sequence analysis. Forensic scientists are expected to appear at court hearings to present their findings. They are usually employed in crime labs of city and state government agencies. Geneticists experimenting with DNA techniques also work for scientific and research organizations, pharmaceutical industries, and college and university labs. Students wishing to pursue a career as a forensic scientist should have at least a bachelor's degree in chemistry, biology, or physics, and preferably some experience working in a laboratory.

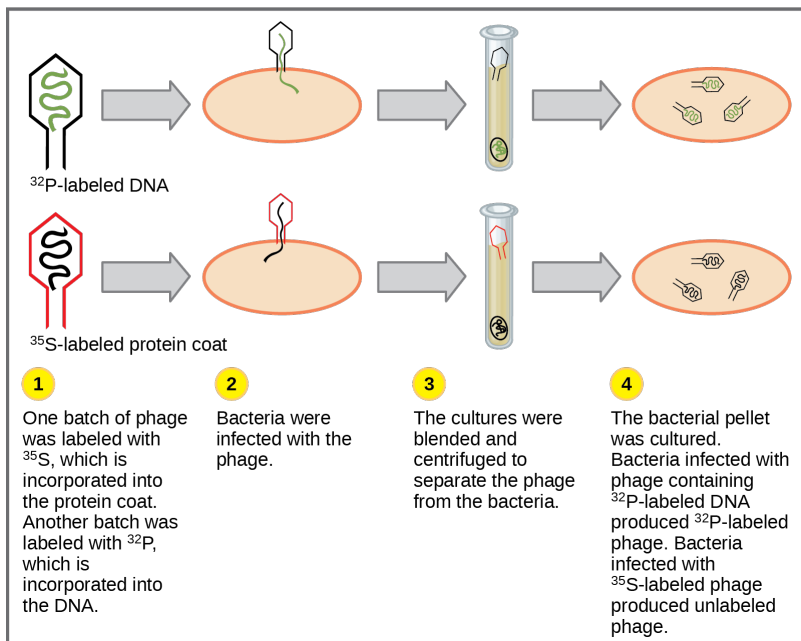
Although the experiments of Avery, McCarty and McLeod had demonstrated that DNA was the informational component transferred during transformation, DNA was still considered to be too simple a molecule to carry biological information. Proteins, with their 20 different amino acids, were regarded as more likely candidates. The decisive experiment, conducted by Martha Chase and Alfred Hershey in 1952, provided confirmatory evidence that DNA was indeed the genetic material and not proteins. Chase and Hershey were studying a **bacteriophage**—a virus that infects bacteria. Viruses typically have a simple structure: a protein coat, called the capsid, and a nucleic acid core that

contains the genetic material (either DNA or RNA). The bacteriophage infects the host bacterial cell by attaching to its surface, and then it injects its nucleic acids inside the cell. The phage DNA makes multiple copies of itself using the host machinery, and eventually the host cell bursts, releasing a large number of bacteriophages. Hershey and Chase selected radioactive elements that would specifically distinguish the protein from the DNA in infected cells. They labeled one batch of phage with radioactive sulfur, ^{35}S , to label the protein coat. Another batch of phage were labeled with radioactive phosphorus, ^{32}P . Because phosphorus is found in DNA, but not protein, the DNA and not the protein would be tagged with radioactive phosphorus. Likewise, sulfur is absent from DNA, but present in several amino acids such as methionine and cysteine.

Each batch of phage was allowed to infect the cells separately. After infection, the phage bacterial suspension was put in a blender, which caused the phage coat to detach from the host cell. Cells exposed long enough for infection to occur were then examined to see which of the two radioactive molecules had entered the cell. The phage and bacterial suspension was spun down in a centrifuge. The heavier bacterial cells settled down and formed a pellet, whereas the lighter phage particles stayed in the supernatant. In the tube that contained phage labeled with ^{35}S , the supernatant contained the

radioactively labeled phage, whereas no radioactivity was detected in the pellet. In the tube that contained the phage labeled with ^{32}P , the radioactivity was detected in the pellet that contained the heavier bacterial cells, and no radioactivity was detected in the supernatant. Hershey and Chase concluded that it was the phage DNA that was injected into the cell and carried information to produce more phage particles, thus providing evidence that DNA was the genetic material and not proteins ([\[link\]](#)).

In Hershey and Chase's experiments, bacteria were infected with phage radiolabeled with either ^{35}S , which labels protein, or ^{32}P , which labels DNA. Only ^{32}P entered the bacterial cells, indicating that DNA is the genetic material.



Around this same time, Austrian biochemist Erwin Chargaff examined the content of DNA in different species and found that the amounts of adenine, thymine, guanine, and cytosine were not found in equal quantities, and that relative concentrations of the four nucleotide bases varied from species to species, but not within tissues of the same individual or between individuals of the same species. He also discovered something unexpected: That the amount of adenine equaled the amount of thymine, and the amount of cytosine equaled the amount of guanine (that is, $A = T$ and $G = C$). Different species had equal amounts of *purines* ($A + G$) and **pyrimidines** ($T + C$), but different ratios of $A + T$ to $G + C$. These observations became known as **Chargaff's rules**. Chargaff's findings proved immensely useful when Watson and Crick were getting ready to propose their DNA double helix model! You can see after reading the past few pages how science builds upon previous discoveries, sometimes in a slow and laborious process.

Section Summary

DNA was first isolated from white blood cells by Friedrich Miescher, who called it nuclein because it was isolated from nuclei. Frederick Griffith's experiments with strains of *Streptococcus pneumoniae* provided the first hint that DNA may be the transforming principle. Avery, MacLeod, and

McCarty showed that DNA is required for the transformation of bacteria. Later experiments by Hershey and Chase using bacteriophage T2 proved that DNA is the genetic material. Chargaff found that the ratio of $A = T$ and $C = G$, and that the percentage content of A, T, G, and C is different for different species.

Review Questions

If DNA of a particular species was analyzed and it was found that it contains 27 percent A, what would be the percentage of C?

1. 27 percent
2. 30 percent
3. 23 percent
4. 54 percent

C

The experiments by Hershey and Chase helped confirm that DNA was the hereditary material on the basis of the finding that:

1. radioactive phage were found in the pellet
2. radioactive cells were found in the

supernatant

3. radioactive sulfur was found inside the cell
4. radioactive phosphorus was found in the cell

D

Bacterial transformation is a major concern in many medical settings. Why might health care providers be concerned?

1. Pathogenic bacteria could introduce disease-causing genes in non-pathogenic bacteria.
2. Antibiotic resistance genes could be introduced to new bacteria to create “superbugs.”
3. Bacteriophages could spread DNA encoding toxins to new bacteria.
4. All of the above.

D

Critical Thinking Questions

Explain Griffith's transformation experiments.
What did he conclude from them?

Live R cells acquired genetic information from the heat-killed S cells that “transformed” the R cells into S cells.

Why were radioactive sulfur and phosphorous used to label bacteriophage in Hershey and Chase's experiments?

Sulfur is an element found in proteins and phosphorus is a component of nucleic acids.

When Chargaff was performing his experiments, the tetranucleotide hypothesis, which stated that DNA was composed of GACT nucleotide repeats, was the most widely accepted view of DNA's composition. How did Chargaff disprove this hypothesis?

If the tetranucleotide hypothesis were true, then DNA would have to contain equal amounts of all four nucleotides ($A = T = G = C$). However, Chargaff demonstrated that $A = T$ and $G = C$, but that the four nucleotides are not present in equal amounts.

Glossary

transformation

process in which external DNA is taken up by
a cell

DNA Structure and Sequencing

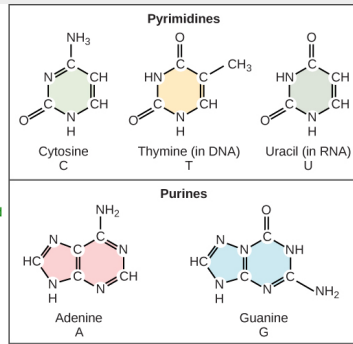
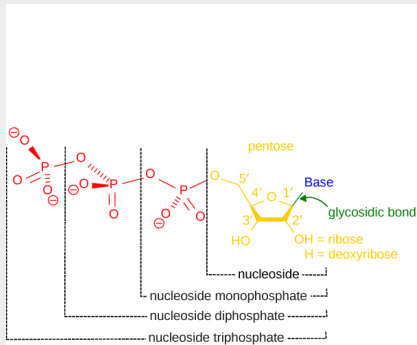
By the end of this section, you will be able to do the following:

- Describe the structure of DNA
- Explain the Sanger method of DNA sequencing
- Discuss the similarities and differences between eukaryotic and prokaryotic DNA

The building blocks of DNA are nucleotides. The important components of the nucleotide are a nitrogenous (nitrogen-bearing) base, a 5-carbon sugar (pentose), and a phosphate group ([\[link\]](#)). The nucleotide is named depending on the nitrogenous base. The nitrogenous base can be a purine such as adenine (A) and guanine (G), or a pyrimidine such as cytosine (C) and thymine (T).

Visual Connection

The purines have a double ring structure with a six-membered ring fused to a five-membered ring. Pyrimidines are smaller in size; they have a single six-membered ring structure.



The images above illustrate the five bases of DNA and RNA. Examine the images and explain why these are called “nitrogenous bases.” How are the purines different from the pyrimidines? How is one purine or pyrimidine different from another, e.g., adenine from guanine? How is a nucleoside different from a nucleotide?

The purines have a double ring structure with a six-membered ring fused to a five-membered ring. Pyrimidines are smaller in size; they have a single six-membered ring structure.

The sugar is deoxyribose in DNA and ribose in RNA. The carbon atoms of the five-carbon sugar are numbered 1', 2', 3', 4', and 5' (1' is read as “one prime”). The phosphate, which makes DNA and RNA acidic, is connected to the 5' carbon of the sugar by the formation of an ester linkage between phosphoric acid and the 5'-OH group (an ester is an acid + an alcohol). In DNA nucleotides, the 3' carbon of the sugar deoxyribose is attached to a hydroxyl (OH) group. In RNA nucleotides, the 2' carbon of the sugar ribose also contains a hydroxyl group. The base is attached to the 1' carbon of the sugar.

The nucleotides combine with each other to produce phosphodiester bonds. The phosphate residue attached to the 5' carbon of the sugar of one nucleotide forms a second ester linkage with the hydroxyl group of the 3' carbon of the sugar of the next nucleotide, thereby forming a 5'-3' phosphodiester bond. In a polynucleotide, one end of the chain has a free 5' phosphate, and the other end has a free 3'-OH. These are called the 5' and 3' ends of the chain.

In the 1950s, Francis Crick and James Watson worked together to determine the structure of DNA at the University of Cambridge, England. Other scientists like Linus Pauling and Maurice Wilkins were also actively exploring this field. Pauling previously had discovered the secondary structure of proteins using X-ray crystallography. In Wilkins' lab, researcher Rosalind Franklin was using X-ray diffraction methods to understand the structure of DNA. Watson and Crick were able to piece together the puzzle of the DNA molecule on the basis of Franklin's data because Crick had also studied X-ray diffraction ([\[link\]](#)). In 1962, James Watson, Francis Crick, and Maurice Wilkins were awarded the Nobel Prize in Medicine. Unfortunately, by then Franklin had died, and Nobel prizes are not awarded posthumously.

The work of pioneering scientists (a) James Watson, Francis Crick, and Maclyn McCarty led to our present day understanding of DNA. Scientist

Rosalind Franklin discovered (b) the X-ray diffraction pattern of DNA, which helped to elucidate its double-helix structure. (credit a: modification of work by Marjorie McCarty, Public Library of Science)



(a)

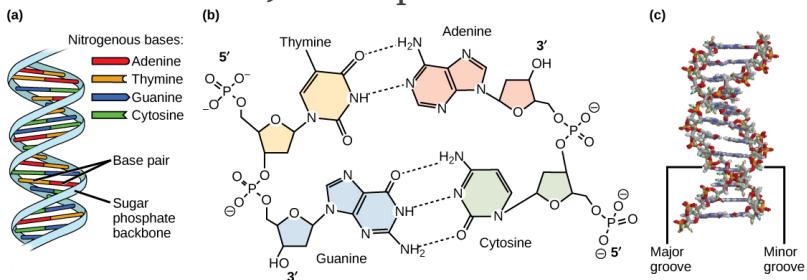


(b)

Watson and Crick proposed that DNA is made up of two strands that are twisted around each other to form a right-handed helix. Base pairing takes place between a purine and pyrimidine on opposite strands, so that A pairs with T, and G pairs with C (suggested by Chargaff's Rules). Thus, adenine and thymine are complementary base pairs, and cytosine and guanine are also complementary base pairs. The base pairs are stabilized by hydrogen bonds: *adenine and thymine form two hydrogen bonds and cytosine and guanine form three hydrogen bonds*. The two strands are anti-parallel in nature; that is, the 3' end of one strand faces the 5' end of the other strand. The sugar and phosphate of the nucleotides form the backbone of the structure, whereas the nitrogenous bases are stacked inside, like the rungs of a ladder.

Each base pair is separated from the next base pair by a distance of 0.34 nm, and each turn of the helix measures 3.4 nm. Therefore, 10 base pairs are present per turn of the helix. The diameter of the DNA double-helix is 2 nm, and it is uniform throughout. Only the pairing between a purine and pyrimidine and the antiparallel orientation of the two DNA strands can explain the uniform diameter. The twisting of the two strands around each other results in the formation of uniformly spaced major and minor grooves ([\[link\]](#)).

DNA has (a) a double helix structure and (b) phosphodiester bonds; the dotted lines between Thymine and Adenine and Guanine and Cytosine represent hydrogen bonds. The (c) major and minor grooves are binding sites for DNA binding proteins during processes such as transcription (the copying of RNA from DNA) and replication.



In Frederick Sanger's dideoxy chain termination method, dye-labeled dideoxynucleotides are used to generate DNA fragments that terminate at different points. The DNA is separated by capillary electrophoresis (not defined) on the basis of size, and from the order of fragments formed, the DNA sequence can be read. The DNA sequence readout is

shown on an electropherogram (not defined) that is generated by a laser scanner. DNA can be separated on the basis of size using gel electrophoresis. (credit: James Jacob, Tompkins Cortland Community College) These figures illustrate the compaction of the eukaryotic chromosome.

DNA Sequencing Techniques

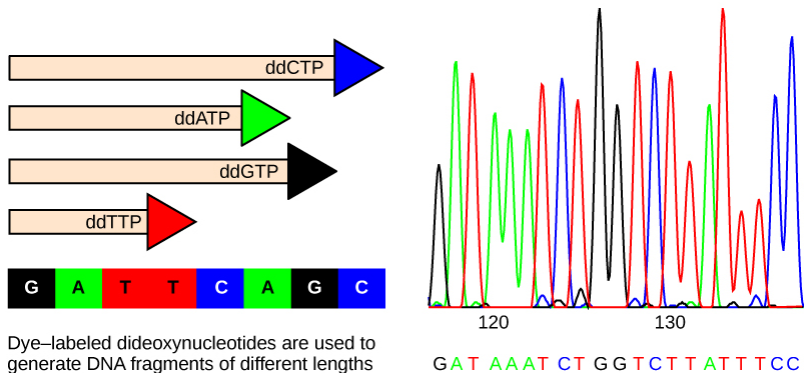
Until the 1990s, the sequencing of DNA (reading the sequence of DNA) was a relatively expensive and long process. Using radiolabeled nucleotides also compounded the problem through safety concerns. With currently available technology and automated machines, the process is cheaper, safer, and can be completed in a matter of hours. Fred Sanger developed the sequencing method used for the human genome sequencing project, which is widely used today ([\[link\]](#)).

Link to Learning

Visit [this site](#) to watch a video explaining the DNA sequence-reading technique that resulted from Sanger's work.

The sequencing method is known as the dideoxy chain termination method. The method is based on

the use of chain terminators, the dideoxynucleotides (ddNTPs). The ddNTPs differ from the deoxynucleotides by the lack of a free 3' OH group on the five-carbon sugar. If a ddNTP is added to a growing DNA strand, the chain cannot be extended any further because the free 3' OH group needed to add another nucleotide is not available. By using a predetermined ratio of deoxyribonucleotides to dideoxynucleotides, it is possible to generate DNA fragments of different sizes.



The DNA sample to be sequenced is denatured (separated into two strands by heating it to high temperatures). The DNA is divided into four tubes in which a primer, DNA polymerase, and all four nucleoside triphosphates (A, T, G, and C) are added. In addition, limited quantities of one of the four dideoxynucleoside triphosphates (ddCTP, ddATP, ddGTP, and ddTTP) are added to each tube respectively. The tubes are labeled as A, T, G, and C according to the ddNTP added. For detection purposes, each of the four dideoxynucleotides carries a different fluorescent label. Chain

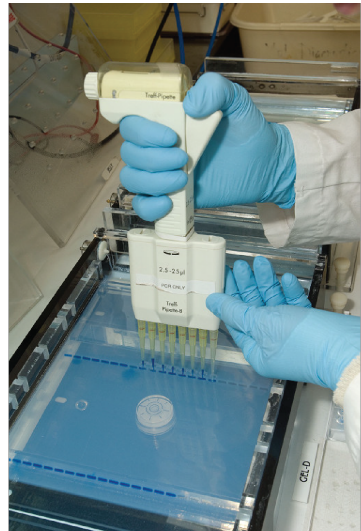
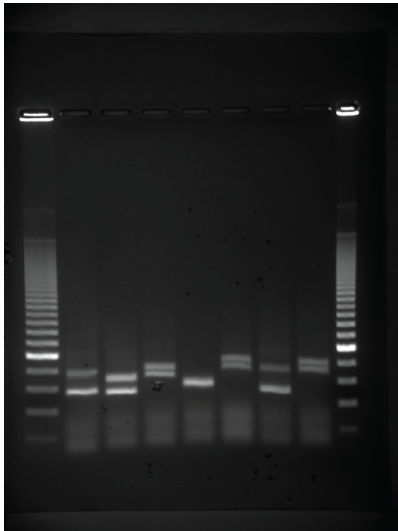
elongation continues until a fluorescent dideoxy nucleotide is incorporated, after which no further elongation takes place. After the reaction is over, electrophoresis is performed. Even a difference in length of a single base can be detected. The sequence is read from a laser scanner that detects the fluorescent marker of each fragment. For his work on DNA sequencing, Sanger received a Nobel Prize in Chemistry in 1980.

Link to Learning

Sanger's genome sequencing has led to a race to sequence human genomes at rapid speed and low cost, often referred to as the \$1000-in-one-day sequence. Learn more by selecting the *Sequencing at Speed* animation [here](#).

Gel electrophoresis is a technique used to separate DNA fragments of different sizes. Usually the gel is made of a chemical called *agarose* (a polysaccharide polymer extracted from seaweed that is high in galactose residues). Agarose powder is added to a buffer and heated. After cooling, the gel solution is poured into a casting tray. Once the gel has solidified, the DNA is loaded on the gel and electric current is applied. The DNA has a net negative charge and moves from the negative electrode

toward the positive electrode. The electric current is applied for sufficient time to let the DNA separate according to size; the smallest fragments will be farthest from the well (where the DNA was loaded), and the heavier molecular weight fragments will be closest to the well. Once the DNA is separated, the gel is stained with a DNA-specific dye for viewing it ([\[link\]](#)).



Evolution Connection

Neanderthal Genome: How Are We Related?

The first draft sequence of the Neanderthal genome was recently published by Richard E. Green et al. in 2010.[\[footnote\]](#) Neanderthals are the closest ancestors of present-day humans. They were known to have lived in Europe and Western Asia (and now, perhaps, in Northern Africa) before they

disappeared from fossil records approximately 30,000 years ago. Green's team studied almost 40,000-year-old fossil remains that were selected from sites across the world. Extremely sophisticated means of sample preparation and DNA sequencing were employed because of the fragile nature of the bones and heavy microbial contamination. In their study, the scientists were able to sequence some four billion base pairs. The Neanderthal sequence was compared with that of present-day humans from across the world. After comparing the sequences, the researchers found that the Neanderthal genome had 2 to 3 percent greater similarity to people living outside Africa than to people in Africa. While current theories have suggested that all present-day humans can be traced to a small ancestral population in Africa, the data from the Neanderthal genome suggest some interbreeding between Neanderthals and early modern humans.

Richard E. Green et al., "A Draft Sequence of the Neandertal Genome," *Science* 328 (2010): 710-22. Green and his colleagues also discovered DNA segments among people in Europe and Asia that are more similar to Neanderthal sequences than to other contemporary human sequences. Another interesting observation was that Neanderthals are as closely related to people from Papua New Guinea as to those from China or France. This is surprising because Neanderthal fossil remains have been located only in Europe and West Asia. Most

likely, genetic exchange took place between Neanderthals and modern humans as modern humans emerged out of Africa, before the divergence of Europeans, East Asians, and Papua New Guineans.

Several genes seem to have undergone changes from Neanderthals during the evolution of present-day humans. These genes are involved in cranial structure, metabolism, skin morphology, and cognitive development. One of the genes that is of particular interest is *RUNX2*, which is different in modern day humans and Neanderthals. This gene is responsible for the prominent frontal bone, bell-shaped rib cage, and dental differences seen in Neanderthals. It is speculated that an evolutionary change in *RUNX2* was important in the origin of modern-day humans, and this affected the cranium and the upper body.

Link to Learning

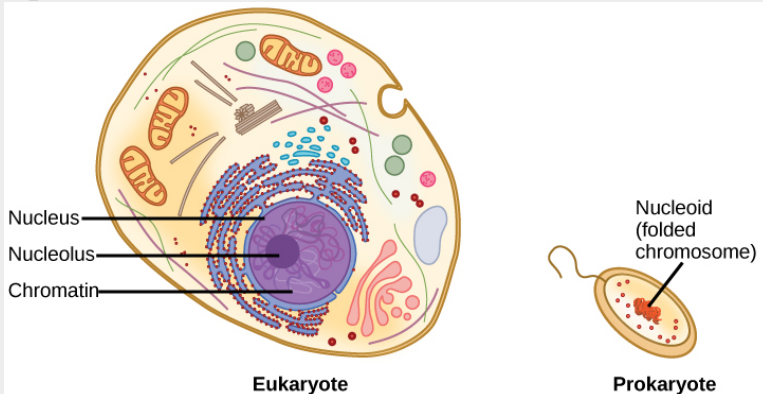
Watch [Svante Pääbo's talk](#) explaining the Neanderthal genome research at the 2011 annual TED (Technology, Entertainment, Design) conference.

DNA Packaging in Cells

Prokaryotes are much simpler than eukaryotes in many of their features ([\[link\]](#)). Most prokaryotes contain a single, circular chromosome that is found in an area of the cytoplasm called the *nucleoid region*.

Visual Connection

A eukaryote contains a well-defined nucleus, whereas in prokaryotes, the chromosome lies in the cytoplasm in an area called the nucleoid.



In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them occur together?

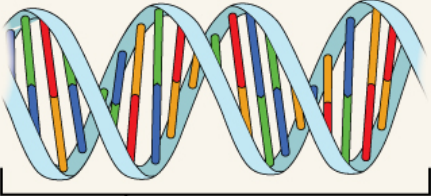
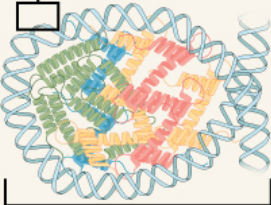
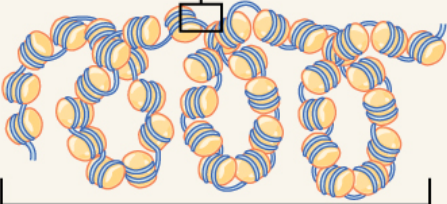
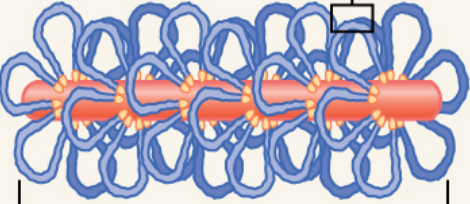
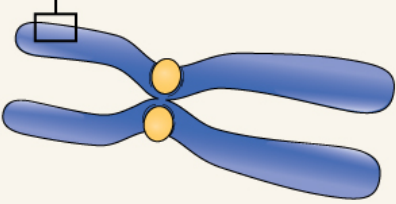
The size of the genome in one of the most well-

studied prokaryotes, *E.coli*, is 4.6 million base pairs (approximately 1.1 mm, if cut and stretched out). So how does this fit inside a small bacterial cell? The DNA is twisted by what is known as supercoiling. Supercoiling suggests that DNA is either “under-wound” (less than one turn of the helix per 10 base pairs) or “over-wound” (more than 1 turn per 10 base pairs) from its normal relaxed state. Some proteins are known to be involved in the supercoiling; other proteins and enzymes such as DNA gyrase help in maintaining the supercoiled structure.

Eukaryotes, whose chromosomes each consist of a linear DNA molecule, employ a different type of packing strategy to fit their DNA inside the nucleus ([\[link\]](#)). At the most basic level, DNA is wrapped around proteins known as **histones** to form structures called nucleosomes. The histones are evolutionarily conserved proteins that are rich in basic amino acids and form an octamer composed of two molecules of each of four different histones. The DNA (remember, it is negatively charged because of the phosphate groups) is wrapped tightly around the histone core. This nucleosome is linked to the next one with the help of a *linker DNA*. This is also known as the “beads on a string” structure. With the help of a fifth histone, a string of nucleosomes is further compacted into a 30-nm fiber, which is the diameter of the structure. Metaphase chromosomes are even further condensed by association with

scaffolding proteins. At the metaphase stage, the chromosomes are at their most compact, approximately 700 nm in width.

In interphase, eukaryotic chromosomes have two distinct regions that can be distinguished by staining. The tightly packaged region is known as heterochromatin, and the less dense region is known as euchromatin. Heterochromatin usually contains genes that are not expressed, and is found in the regions of the centromere and telomeres. The euchromatin usually contains genes that are transcribed, with DNA packaged around nucleosomes but not further compacted.

Organization of Eukaryotic Chromosomes	
DNA double helix	
DNA wrapped around histone	
Nucleosomes coiled into a chromatin fiber	
Further condensation of chromatin	
Duplicated chromosome	

Section Summary

The currently accepted model of the double-helix structure of DNA was proposed by Watson and Crick. Some of the salient features are that the two strands that make up the double helix have complementary base sequences and anti-parallel orientations. Alternating deoxyribose sugars and phosphates form the backbone of the structure, and the nitrogenous bases are stacked like rungs inside. The diameter of the double helix, 2 nm, is uniform throughout. A purine always pairs with a pyrimidine; A pairs with T, and G pairs with C. One turn of the helix has 10 base pairs. Prokaryotes are much simpler than eukaryotes in many of their features. Most prokaryotes contain a single, circular chromosome. In general, eukaryotic chromosomes contain a linear DNA molecule packaged into nucleosomes, and have two distinct regions that can be distinguished by staining, reflecting different states of packaging and compaction.

Visual Connection Questions

[\[link\]](#) In eukaryotic cells, DNA and RNA synthesis occur in a separate compartment from protein synthesis. In prokaryotic cells, both processes occur together. What advantages might there be to separating the processes? What advantages might there be to having them

occur together?

[\[link\]](#) Compartmentalization enables a eukaryotic cell to divide processes into discrete steps so it can build more complex protein and RNA products. But there is an advantage to having a single compartment as well: RNA and protein synthesis occurs much more quickly in a prokaryotic cell.

Review Questions

DNA double helix does not have which of the following?

1. antiparallel configuration
2. complementary base pairing
3. major and minor grooves
4. uracil

D

In eukaryotes, what is the DNA wrapped around?

1. single-stranded binding proteins
 2. sliding clamp
 3. polymerase
 4. histones
-

D

Critical Thinking Questions

Provide a brief summary of the Sanger sequencing method.

The template DNA strand is mixed with a DNA polymerase, a primer, the 4 deoxynucleotides, and a limiting concentration of 4 dideoxynucleotides. DNA polymerase synthesizes a strand complementary to the template. Incorporation of ddNTPs at different locations results in DNA fragments that have terminated at every possible base in the template. These fragments are separated by gel electrophoresis and visualized by a laser detector to determine the sequence of bases.

Describe the structure and complementary base

pairing of DNA.

DNA has two strands in anti-parallel orientation. The sugar-phosphate linkages form a backbone on the outside, and the bases are paired on the inside: A with T, and G with C, like rungs on a spiral ladder.

Prokaryotes have a single circular chromosome while eukaryotes have linear chromosomes. Describe one advantage and one disadvantage to the eukaryotic genome packaging compared to the prokaryotes.

Advantage: The linear arrangement of the eukaryotic chromosome allows more DNA to be packed by tightly winding it around histones. More genetic material means that the organism can encode more information into a single cell. This eventually allowed some eukaryotes to develop into multicellular organisms with cell specialization. Disadvantage: Maintaining more genetic material requires more energy, and introduces the possibility for more errors (more complexity).

Glossary

electrophoresis

technique used to separate DNA fragments
according to size

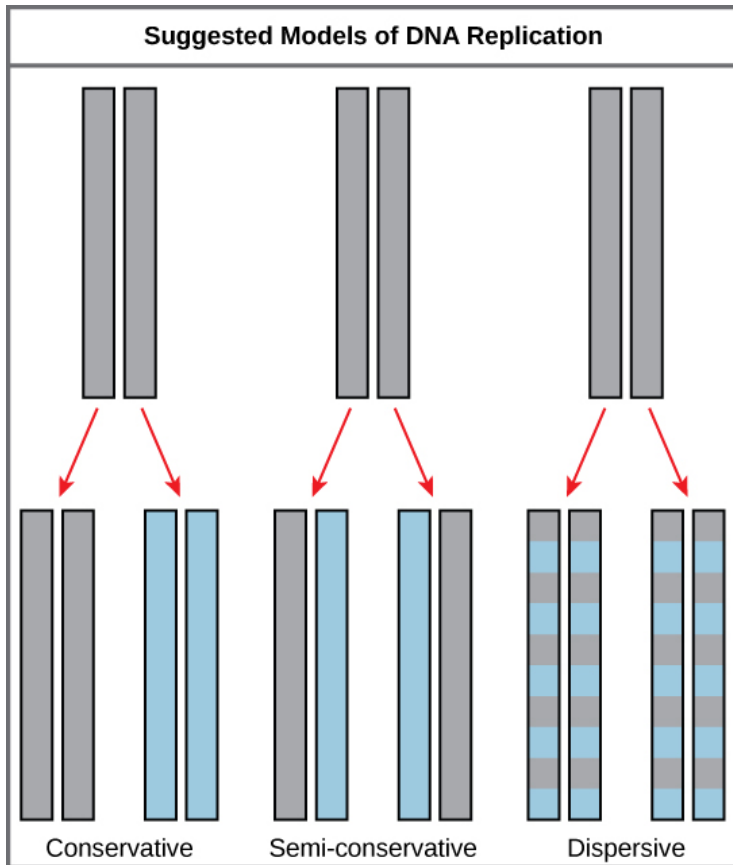
Basics of DNA Replication

By the end of this section, you will be able to do the following:

- Explain how the structure of DNA reveals the replication process
- Describe the Meselson and Stahl experiments

The elucidation of the structure of the double helix provided a hint as to how DNA divides and makes copies of itself. In their 1953 paper, Watson and Crick penned an incredible understatement: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." With specific base pairs, the sequence of one DNA strand can be predicted from its complement. The double-helix model suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. What was not clear was how the replication took place. There were three models suggested ([\[link\]](#)): *conservative*, *semi-conservative*, and *dispersive*.

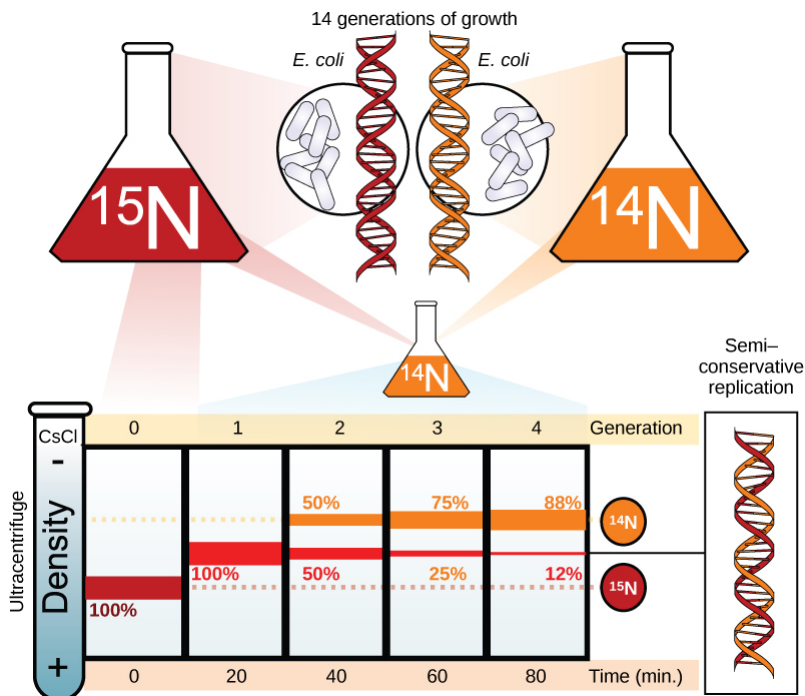
The three suggested models of DNA replication. Gray indicates the original DNA strands, and blue indicates newly synthesized DNA.



In conservative replication, the parental DNA remains together, and the newly formed daughter strands are together. The semi-conservative method suggests that each of the two parental DNA strands acts as a template for new DNA to be synthesized; after replication, each double-stranded DNA includes one parental or “old” strand and one “new” strand. In the dispersive model, both copies of DNA have double-stranded segments of parental DNA and newly synthesized DNA interspersed.

Meselson and Stahl were interested in understanding how DNA replicates. They grew *E. coli* for several generations in a medium containing a “heavy” isotope of nitrogen (^{15}N), which gets incorporated into nitrogenous bases, and eventually into the DNA ([\[link\]](#)).

Meselson and Stahl experimented with *E. coli* grown first in heavy nitrogen (^{15}N) then in ^{14}N . DNA grown in ^{15}N (red band) is heavier than DNA grown in ^{14}N (orange band), and sediments to a lower level in cesium chloride solution in an ultracentrifuge. When DNA grown in ^{15}N is switched to media containing ^{14}N , after one round of cell division the DNA sediments halfway between the ^{15}N and ^{14}N levels, indicating that it now contains fifty percent ^{14}N . In subsequent cell divisions, an increasing amount of DNA contains ^{14}N only. These data support the semi-conservative replication model. (credit: modification of work by Mariana Ruiz Villareal)



The *E. coli* culture was then placed into medium containing ^{14}N and allowed to grow for several generations. After each of the first few generations, the cells were harvested and the DNA was isolated, then centrifuged at high speeds in an ultracentrifuge. During the centrifugation, the DNA was loaded into a *gradient* (typically a solution of salt such as cesium chloride or sucrose) and spun at high speeds of 50,000 to 60,000 rpm. Under these circumstances, the DNA will form a band according to its *buoyant density*: the density within the gradient at which it floats. DNA grown in ^{15}N will form a band at a higher density position (i.e., farther down the centrifuge tube) than that grown in ^{14}N . Meselson and Stahl noted that after one generation

of growth in ^{14}N after they had been shifted from ^{15}N , the single band observed was intermediate in position in between DNA of cells grown exclusively in ^{15}N and ^{14}N . This suggested either a semi-conservative or dispersive mode of replication. The DNA harvested from cells grown for two generations in ^{14}N formed two bands: one DNA band was at the intermediate position between ^{15}N and ^{14}N , and the other corresponded to the band of ^{14}N DNA. These results could only be explained if DNA replicates in a semi-conservative manner. And for this reason, therefore, the other two models were ruled out.

During DNA replication, each of the two strands that make up the double helix serves as a template from which new strands are copied. The new strands will be complementary to the parental or “old” strands. When two daughter DNA copies are formed, they have the same sequence and are divided equally into the two daughter cells.

Link to Learning

View [this video](#) on DNA replication.

Section Summary

During cell division, each daughter cell receives a copy of each molecule of DNA by a process known as DNA replication. The single chromosome of a prokaryote or each chromosome of a eukaryote consists of a single continuous double helix. The model for DNA replication suggests that the two strands of the double helix separate during replication, and each strand serves as a template from which the new complementary strand is copied. In the conservative model of replication, the parental DNA is conserved, and the daughter DNA is newly synthesized. The semi-conservative model suggests that each of the two parental DNA strands acts as template for new DNA to be synthesized; after replication, each double-stranded DNA retains the parental or “old” strand and one “new” strand. The dispersive model suggested that the two copies of the DNA would have segments of parental DNA and newly synthesized DNA. The Meselson and Stahl experiment supported the semi-conservative model of replication, in which an entire replicated chromosome consists of one parental strand and one newly synthesized strand of DNA.

Review Questions

Meselson and Stahl's experiments proved that DNA replicates by which mode?

1. conservative
 2. semi-conservative
 3. dispersive
 4. none of the above
-

B

If the sequence of the 5'-3' strand is AATGCTAC, then the complementary sequence has the following sequence:

1. 3'-AATGCTAC-5'
 2. 3'-CATCGTAA-5'
 3. 3'-TTACGATG-5'
 4. 3'-GTAGCATT-5'
-

C

How did Meselson and Stahl support Watson and Crick's double-helix model?

1. They demonstrated that each strand serves as a template for synthesizing a new strand of DNA.
2. They showed that the DNA strands break and recombine without losing genetic material.
3. They proved that DNA maintains a double-

helix structure while undergoing semi-conservative replication.

4. They demonstrated that conservative replication maintains the complementary base pairing of each DNA helix.

A

Critical Thinking Questions

How did the scientific community learn that DNA replication takes place in a semi-conservative fashion?

Meselson's experiments with *E. coli* grown in ^{15}N deduced this finding.

Imagine the Meselson and Stahl experiments had supported conservative replication instead of semi-conservative replication. What results would you predict to observe after two rounds of replication? Be specific regarding percent distributions of DNA incorporating ^{15}N and ^{14}N in the gradient.

Following two rounds of conservative replication, two bands would be detected after ultracentrifugation. A lower (heavier) band would be at the ^{15}N density, and would comprise 25% of the total DNA. A second, higher (lighter) band would be at the ^{14}N density, and would contain 75% of the total DNA.

DNA Replication in Prokaryotes

By the end of this section, you will be able to do the following:

- Explain the process of DNA replication in prokaryotes
- Discuss the role of different enzymes and proteins in supporting this process

DNA replication has been well studied in prokaryotes primarily because of the small size of the genome and because of the large variety of mutants that are available. *E. coli* has 4.6 million base pairs in a single circular chromosome and all of it gets replicated in approximately 42 minutes, starting from a single site along the chromosome and proceeding around the circle in both directions. This means that approximately 1000 nucleotides are added per second. Thus, the process is quite rapid and occurs without many mistakes.

DNA replication employs a large number of structural proteins and enzymes, each of which plays a critical role during the process. One of the key players is the enzyme **DNA polymerase**, also known as DNA pol, which adds nucleotides one-by-one to the growing DNA chain that is complementary to the template strand. The addition of nucleotides requires energy; this energy is obtained from the nucleoside triphosphates ATP, GTP, TTP and CTP. Like ATP, the other **NTPs**

(nucleoside triphosphates) are high-energy molecules that can serve both as the source of DNA nucleotides and the source of energy to drive the polymerization. When the bond between the phosphates is “broken,” the energy released is used to form the phosphodiester bond between the incoming nucleotide and the growing chain. In prokaryotes, three main types of polymerases are known: DNA pol I, DNA pol II, and DNA pol III. It is now known that DNA pol III is the enzyme required for DNA synthesis; DNA pol I is an important accessory enzyme in DNA replication, and along with DNA pol II, is primarily required for repair.

How does the replication machinery know where to begin? It turns out that there are specific nucleotide sequences called *origins of replication* where replication begins. In *E. coli*, which has a single origin of replication on its one chromosome (as do most prokaryotes), this origin of replication is approximately 245 base pairs long and is rich in AT sequences. The origin of replication is recognized by certain proteins that bind to this site. An enzyme called *helicase* unwinds the DNA by breaking the hydrogen bonds between the nitrogenous base pairs. ATP hydrolysis is required for this process. As the DNA opens up, Y-shaped structures called *replication forks* are formed. Two replication forks are formed at the origin of replication and these get extended bi-directionally as replication proceeds. **Single-strand binding proteins** coat the single strands of

DNA near the replication fork to prevent the single-stranded DNA from winding back into a double helix.

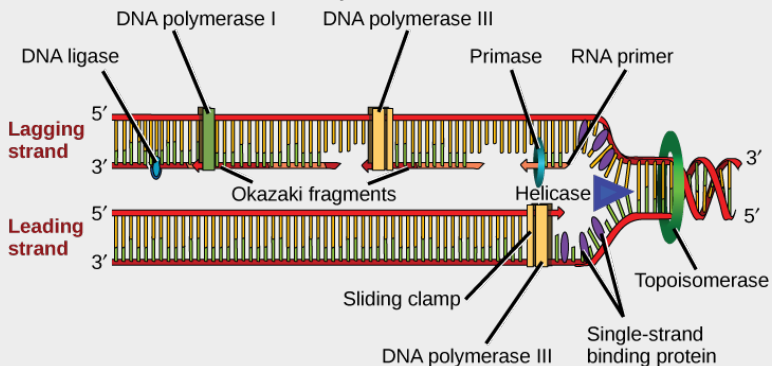
DNA polymerase has two important restrictions: it is able to add nucleotides only in the 5' to 3' direction (a new DNA strand can be only extended in this direction). It also requires a free 3'-OH group to which it can add nucleotides by forming a phosphodiester bond between the 3'-OH end and the 5' phosphate of the next nucleotide. This essentially means that it cannot add nucleotides if a free 3'-OH group is not available. Then how does it add the first nucleotide? The problem is solved with the help of a primer that provides the free 3'-OH end.

Another enzyme, RNA **primase**, synthesizes an RNA segment that is about five to ten nucleotides long and complementary to the template DNA. Because this sequence primes the DNA synthesis, it is appropriately called the **primer**. DNA polymerase can now extend this RNA primer, adding nucleotides one-by-one that are complementary to the template strand ([\[link\]](#)).

Visual Connection

A replication fork is formed when helicase separates the DNA strands at the origin of replication. The DNA tends to become more highly coiled ahead of the replication fork. Topoisomerase

breaks and reforms DNA's phosphate backbone ahead of the replication fork, thereby relieving the pressure that results from this "supercoiling." Single-strand binding proteins bind to the single-stranded DNA to prevent the helix from re-forming. Primase synthesizes an RNA primer. DNA polymerase III uses this primer to synthesize the daughter DNA strand. On the leading strand, DNA is synthesized continuously, whereas on the lagging strand, DNA is synthesized in short stretches called Okazaki fragments. DNA polymerase I replaces the RNA primer with DNA. DNA ligase seals the gaps between the Okazaki fragments, joining the fragments into a single DNA molecule. (credit: modification of work by Mariana Ruiz Villareal)



Question: You isolate a cell strain in which the joining of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

The replication fork moves at the rate of 1000 nucleotides per second. Topoisomerase prevents the over-winding of the DNA double helix ahead of the replication fork as the DNA is opening up; it does so by causing temporary nicks in the DNA helix and then resealing it. Because DNA polymerase can only extend in the 5' to 3' direction, and because the DNA double helix is *antiparallel*, there is a slight problem at the replication fork. *The two template DNA strands have opposing orientations*: one strand is in the 5' to 3' direction and the other is oriented in the 3' to 5' direction. Only one new DNA strand, the one that is complementary to the 3' to 5' parental DNA strand, can be synthesized continuously towards the replication fork. This continuously synthesized strand is known as the **leading strand**. The other strand, complementary to the 5' to 3' parental DNA, is extended away from the replication fork, in small fragments known as **Okazaki fragments**, each requiring a primer to start the synthesis. New primer segments are laid down in the direction of the replication fork, but each pointing away from it. (Okazaki fragments are named after the Japanese scientist who first discovered them. The strand with the Okazaki fragments is known as the **lagging strand**.)

The leading strand can be extended from a single primer, whereas the lagging strand needs a new primer for each of the short Okazaki fragments. The overall direction of the lagging strand will be 3' to

5', and that of the leading strand 5' to 3'. A protein called the **sliding clamp** holds the DNA polymerase in place as it continues to add nucleotides. The sliding clamp is a ring-shaped protein that binds to the DNA and holds the polymerase in place. As synthesis proceeds, the RNA primers are replaced by DNA. The primers are removed by the exonuclease activity of DNA pol I, which uses DNA behind the RNA as its own primer and fills in the gaps left by removal of the RNA nucleotides by the addition of DNA nucleotides. The nicks that remain between the newly synthesized DNA (that replaced the RNA primer) and the previously synthesized DNA are sealed by the enzyme DNA **ligase**, which catalyzes the formation of phosphodiester linkages between the 3'-OH end of one nucleotide and the 5' phosphate end of the other fragment.

Once the chromosome has been completely replicated, the two DNA copies move into two different cells during cell division.

The process of DNA replication can be summarized as follows:

1. DNA unwinds at the origin of replication.
2. Helicase opens up the DNA-forming replication forks; these are extended bidirectionally.
3. Single-strand binding proteins coat the DNA around the replication fork to prevent rewinding of the DNA.

4. Topoisomerase binds at the region ahead of the replication fork to prevent supercoiling.
5. Primase synthesizes RNA primers complementary to the DNA strand.
6. DNA polymerase III starts adding nucleotides to the 3'-OH end of the primer.
7. Elongation of both the lagging and the leading strand continues.
8. RNA primers are removed by exonuclease activity.
9. Gaps are filled by DNA pol I by adding dNTPs.
10. The gap between the two DNA fragments is sealed by DNA ligase, which helps in the formation of phosphodiester bonds.

[\[link\]](#) summarizes the enzymes involved in prokaryotic DNA replication and the functions of each.

Prokaryotic DNA Replication: Enzymes and Their Function

Enzyme/protein	Specific Function
DNA pol I	Removes RNA primer and replaces it with newly synthesized DNA

DNA pol III	Main enzyme that adds nucleotides in the 5'-3' direction
Helicase	Opens the DNA helix by breaking hydrogen bonds between the nitrogenous bases
Ligase	Seals the gaps between the Okazaki fragments to create one continuous DNA strand
Primase	Synthesizes RNA primers needed to start replication
Sliding Clamp	Helps to hold the DNA polymerase in place when nucleotides are being added
Topoisomerase	Helps relieve the strain on DNA when unwinding by causing breaks, and then resealing the DNA
Single-strand binding proteins (SSB)	Binds to single-stranded DNA to prevent DNA from rewinding back.

Link to Learning

Review the full process of DNA replication [here](#).

Section Summary

Replication in prokaryotes starts from a sequence found on the chromosome called the origin of replication—the point at which the DNA opens up. Helicase opens up the DNA double helix, resulting in the formation of the replication fork. Single-strand binding proteins bind to the single-stranded DNA near the replication fork to keep the fork open. Primase synthesizes an RNA primer to initiate synthesis by DNA polymerase, which can add nucleotides only to the 3' end of a previously synthesized primer strand. Both new DNA strands grow according to their respective 5'-3' directions. One strand is synthesized continuously in the direction of the replication fork; this is called the leading strand. The other strand is synthesized in a direction away from the replication fork, in short stretches of DNA known as Okazaki fragments. This strand is known as the lagging strand. Once replication is completed, the RNA primers are replaced by DNA nucleotides and the DNA is sealed with DNA ligase, which creates phosphodiester bonds between the 3'-OH of one end and the 5' phosphate of the other strand.

Visual Connection Questions

[\[link\]](#) You isolate a cell strain in which the joining of Okazaki fragments is impaired and suspect that a mutation has occurred in an enzyme found at the replication fork. Which enzyme is most likely to be mutated?

[\[link\]](#) DNA ligase, as this enzyme joins together Okazaki fragments.

Review Questions

Which of the following components is not involved during the formation of the replication fork?

1. single-strand binding proteins
2. helicase
3. origin of replication
4. ligase

D

Which of the following does the enzyme primase synthesize?

1. DNA primer
 2. RNA primer
 3. Okazaki fragments
 4. phosphodiester linkage
-

B

In which direction does DNA replication take place?

1. 5'-3'
 2. 3'-5'
 3. 5'
 4. 3'
-

A

A scientist randomly mutates the DNA of a bacterium. She then sequences the bacterium's daughter cells, and finds that the daughters have many errors in their replicated DNA. The parent bacterium likely acquired a mutation in which enzyme?

1. DNA ligase
2. DNA pol II
3. Primase
4. DNA pol I

Critical Thinking Questions

DNA replication is bidirectional and discontinuous; explain your understanding of those concepts.

At an origin of replication, two replication forks are formed that are extended in two directions. On the lagging strand, Okazaki fragments are formed in a discontinuous manner.

What are Okazaki fragments and how they are formed?

Short DNA fragments are formed on the lagging strand synthesized in a direction away from the replication fork. These are synthesized by DNA pol.

If the rate of replication in a particular prokaryote is 900 nucleotides per second, how long would it take 1.2 million base pair

genomes to make two copies?

1333 seconds or 22.2 minutes.

Explain the events taking place at the replication fork. If the gene for helicase is mutated, what part of replication will be affected?

At the replication fork, the events taking place are helicase action, binding of single-strand binding proteins, primer synthesis, and synthesis of new strands. If there is a mutated helicase gene, the replication fork will not be extended.

What is the role of a primer in DNA replication? What would happen if you forgot to add a primer in a tube containing the reaction mix for a DNA sequencing reaction?

Primer provides a 3'-OH group for DNA pol to start adding nucleotides. There would be no reaction in the tube without a primer, and no bands would be visible on the electrophoresis.

Quinolone antibiotics treat bacterial infections by blocking the activity of topoisomerase. Why does this treatment work? Explain what occurs at the molecular level.

Bacteria treated with quinolones will no longer be able to replicate their DNA. Topoisomerase relieves the excess DNA supercoiling that occurs ahead of the replication fork as DNA is unwound for replication. If topoisomerase is inhibited, DNA helicase will only be able to unwind the DNA for a short stretch before the supercoiling becomes too overwound for replication to continue.

Glossary

helicase

during replication, this enzyme helps to open up the DNA helix by breaking the hydrogen bonds

lagging strand

during replication, the strand that is replicated in short fragments and away from the replication fork

leading strand

strand that is synthesized continuously in the 5'-3' direction, which is synthesized in the

direction of the replication fork

ligase

enzyme that catalyzes the formation of a phosphodiester linkage between the 3' OH and 5' phosphate ends of the DNA

Okazaki fragment

DNA fragment that is synthesized in short stretches on the lagging strand

primase

enzyme that synthesizes the RNA primer; the primer is needed for DNA pol to start synthesis of a new DNA strand

primer

short stretch of nucleotides that is required to initiate replication; in the case of replication, the primer has RNA nucleotides

replication fork

Y-shaped structure formed during initiation of replication

single-strand binding protein

during replication, protein that binds to the single-stranded DNA; this helps in keeping the two strands of DNA apart so that they may serve as templates

sliding clamp

ring-shaped protein that holds the DNA pol
on the DNA strand

topoisomerase

enzyme that causes underwinding or
overwinding of DNA when DNA replication is
taking place

DNA Replication in Eukaryotes

By the end of this section, you will be able to do the following:

- Discuss the similarities and differences between DNA replication in eukaryotes and prokaryotes
- State the role of telomerase in DNA replication

Eukaryotic genomes are much more complex and larger in size than prokaryotic genomes. Eukaryotes also have a number of different linear chromosomes. The human genome has 3 billion base pairs per haploid set of chromosomes, and 6 billion base pairs are replicated during the S phase of the cell cycle. There are multiple origins of replication on each eukaryotic chromosome; humans can have up to 100,000 origins of replication across the genome. The rate of replication is approximately 100 nucleotides per second, much slower than prokaryotic replication. In yeast, which is a eukaryote, special sequences known as autonomously replicating sequences (ARS) are found on the chromosomes. These are equivalent to the origin of replication in *E. coli*.

The number of DNA polymerases in eukaryotes is much more than in prokaryotes: 14 are known, of which five are known to have major roles during replication and have been well studied. They are known as pol α , pol β , pol γ , pol δ , and pol ϵ .

The essential steps of replication are the same as in prokaryotes. Before replication can start, the DNA has to be made available as a template. Eukaryotic DNA is bound to basic proteins known as histones to form structures called nucleosomes. Histones must be removed and then replaced during the replication process, which helps to account for the lower replication rate in eukaryotes. The chromatin (the complex between DNA and proteins) may undergo some chemical modifications, so that the DNA may be able to slide off the proteins or be accessible to the enzymes of the DNA replication machinery. At the origin of replication, a pre-replication complex is made with other initiator proteins. Helicase and other proteins are then recruited to start the replication process ([link](#)).

Difference between Prokaryotic and Eukaryotic Replication			
Property	Prokaryotes	Eukaryotes	
Origin of replication	Single	Multiple	
Rate of replication	1000 nucleotides/s	50 to 100 nucleotides/s	

DNA polymerase 5 types	5	14
Telomerase	Not present	Present
RNA primer removal	DNA pol I	RNase H
Strand elongation	DNA pol III	Pol α , pol δ , pol ϵ
Sliding clamp	Sliding clamp	PCNA

A helicase using the energy from ATP hydrolysis opens up the DNA helix. Replication forks are formed at each replication origin as the DNA unwinds. The opening of the double helix causes over-winding, or supercoiling, in the DNA ahead of the replication fork. These are resolved with the action of topoisomerases. Primers are formed by the enzyme primase, and using the primer, DNA pol can start synthesis. Three major DNA polymerases are then involved: α , δ and ϵ . DNA pol α adds a short (20 to 30 nucleotides) DNA fragment to the RNA primer on both strands, and then hands off to a second polymerase. While the leading strand is continuously synthesized by the enzyme pol δ , the lagging strand is synthesized by pol ϵ . A sliding clamp protein known as PCNA (proliferating cell nuclear antigen) holds the DNA pol in place so that it does not slide off the DNA. As pol δ runs into the primer RNA on the lagging strand, it displaces it from the DNA template. The displaced primer RNA is then removed by RNase H (AKA flap endonuclease) and replaced with DNA nucleotides.

The Okazaki fragments in the lagging strand are joined after the replacement of the RNA primers with DNA. The gaps that remain are sealed by DNA ligase, which forms the phosphodiester bond.

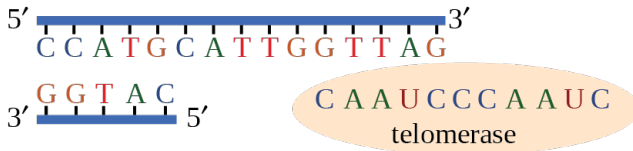
The ends of linear chromosomes are maintained by the action of the telomerase enzyme. Elizabeth Blackburn, 2009 Nobel Laureate, is one of the scientists who discovered how telomerase works. (credit: US Embassy Sweden)Jaskelioff et al., “Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice,” *Nature* 469 (2011): 102-7.

Telomere replication

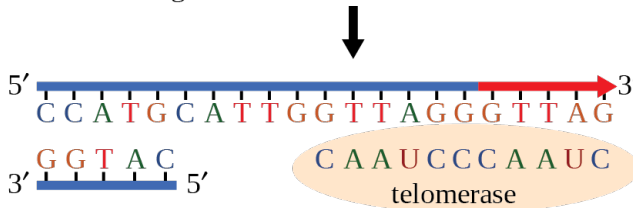
Unlike prokaryotic chromosomes, eukaryotic chromosomes are linear. As you’ve learned, the enzyme DNA pol can add nucleotides only in the 5' to 3' direction. In the leading strand, synthesis continues until the end of the chromosome is reached. On the lagging strand, DNA is synthesized in short stretches, each of which is initiated by a separate primer. When the replication fork reaches the end of the linear chromosome, there is no way to replace the primer on the 5' end of the lagging strand. The DNA at the ends of the chromosome thus remains unpaired, and over time these ends, called **telomeres**, may get progressively shorter as cells continue to divide.

Telomeres comprise repetitive sequences that code

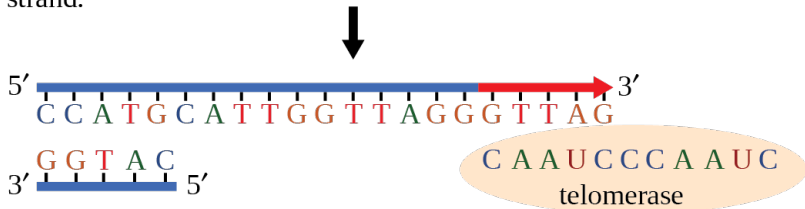
for no particular gene. In humans, a six-base-pair sequence, TTAGGG, is repeated 100 to 1000 times in the telomere regions. In a way, these telomeres protect the genes from getting deleted as cells continue to divide. The telomeres are added to the ends of chromosomes by a separate enzyme, telomerase ([\[link\]](#)), whose discovery helped in the understanding of how these repetitive chromosome ends are maintained. The **telomerase** enzyme contains a catalytic part and a built-in RNA template. It attaches to the end of the chromosome, and DNA nucleotides complementary to the RNA template are added on the 3' end of the DNA strand. Once the 3' end of the lagging strand template is sufficiently elongated, DNA polymerase can add the nucleotides complementary to the ends of the chromosomes. Thus, the ends of the chromosomes are replicated.



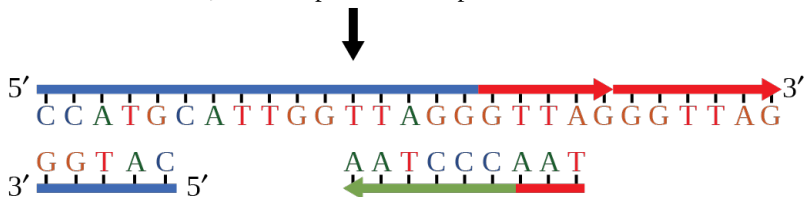
Telomerase has an associated RNA that complements the 3' overhang at the end of the chromosome.



The RNA template is used to synthesize the complementary strand.



Telomerase shifts, and the process is repeated.



Primase and DNA polymerase synthesize the complementary strand.

Telomerase is typically active in germ cells and adult stem cells. It is not active in adult somatic cells. For their discovery of telomerase and its action, Elizabeth Blackburn, Carol W. Greider, and Jack W. Szostak ([\[link\]](#)) received the Nobel Prize for Medicine and Physiology in 2009.



Telomerase and Aging

Cells that undergo cell division continue to have their telomeres shortened because most somatic cells do not make telomerase. This essentially means that telomere shortening is associated with aging. With the advent of modern medicine, preventative health care, and healthier lifestyles, the human life span has increased, and there is an increasing demand for people to look younger and have a better quality of life as they grow older.

In 2010, scientists found that telomerase can reverse some age-related conditions in mice. This may have potential in regenerative medicine.[\[footnote\]](#) Telomerase-deficient mice were used in these studies; these mice have tissue atrophy, stem cell

depletion, organ system failure, and impaired tissue injury responses. Telomerase reactivation in these mice caused extension of telomeres, reduced DNA damage, reversed neurodegeneration, and improved the function of the testes, spleen, and intestines. Thus, telomere reactivation may have potential for treating age-related diseases in humans.

Cancer is characterized by uncontrolled cell division of abnormal cells. The cells accumulate mutations, proliferate uncontrollably, and can migrate to different parts of the body through a process called metastasis. Scientists have observed that cancerous cells have considerably shortened telomeres and that telomerase is active in these cells. Interestingly, only after the telomeres were shortened in the cancer cells did the telomerase become active. If the action of telomerase in these cells can be inhibited by drugs during cancer therapy, then the cancerous cells could potentially be stopped from further division.

Section Summary

Replication in eukaryotes starts at multiple origins of replication. The mechanism is quite similar to that in prokaryotes. A primer is required to initiate synthesis, which is then extended by DNA polymerase as it adds nucleotides one by one to the growing chain. The leading strand is synthesized

continuously, whereas the lagging strand is synthesized in short stretches called Okazaki fragments. The RNA primers are replaced with DNA nucleotides; the DNA Okazaki fragments are linked into one continuous strand by DNA ligase. The ends of the chromosomes pose a problem as the primer RNA at the 5' ends of the DNA cannot be replaced with DNA, and the chromosome is progressively shortened. Telomerase, an enzyme with an inbuilt RNA template, extends the ends by copying the RNA template and extending one strand of the chromosome. DNA polymerase can then fill in the complementary DNA strand using the regular replication enzymes. In this way, the ends of the chromosomes are protected.

Review Questions

The ends of the linear chromosomes are maintained by

1. helicase
2. primase
3. DNA pol
4. telomerase

Which of the following is not a true statement comparing prokaryotic and eukaryotic DNA replication?

1. Both eukaryotic and prokaryotic DNA polymerases build off RNA primers made by primase.
2. Eukaryotic DNA replication requires multiple replication forks, while prokaryotic replication uses a single origin to rapidly replicate the entire genome.
3. DNA replication always occurs in the nucleus.
4. Eukaryotic DNA replication involves more polymerases than prokaryotic replication.

C

Critical Thinking Questions

How do the linear chromosomes in eukaryotes ensure that its ends are replicated completely?

Telomerase has an inbuilt RNA template that extends the 3' end, so primer is synthesized and extended. Thus, the ends are protected.

Glossary

telomerase

enzyme that contains a catalytic part and an inbuilt RNA template; it functions to maintain telomeres at chromosome ends

telomere

DNA at the end of linear chromosomes

DNA Repair

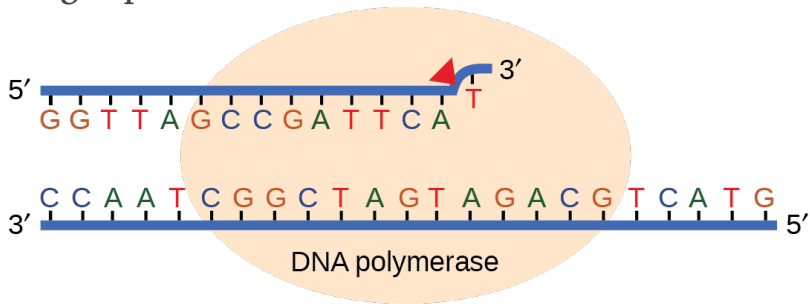
By the end of this section, you will be able to do the following:

- Discuss the different types of mutations in DNA
- Explain DNA repair mechanisms

DNA replication is a highly accurate process, but mistakes can occasionally occur, such as a DNA polymerase inserting a wrong base. Uncorrected mistakes may sometimes lead to serious consequences, such as cancer. Repair mechanisms correct the mistakes. In rare cases, mistakes are not corrected, leading to mutations; in other cases, repair enzymes are themselves mutated or defective.

Most of the mistakes during DNA replication are promptly corrected by the proofreading ability of DNA polymerase itself. ([link](#)). In **proofreading**, the DNA pol reads the newly added base before adding the next one, so a correction can be made. The polymerase checks whether the newly added base has paired correctly with the base in the template strand. If it is the right base, the next nucleotide is added. If an incorrect base has been added, the enzyme makes a cut at the phosphodiester bond and releases the wrong nucleotide. This is performed by the 3' exonuclease action of DNA pol. Once the incorrect nucleotide has been removed, it can be replaced by the correct one.

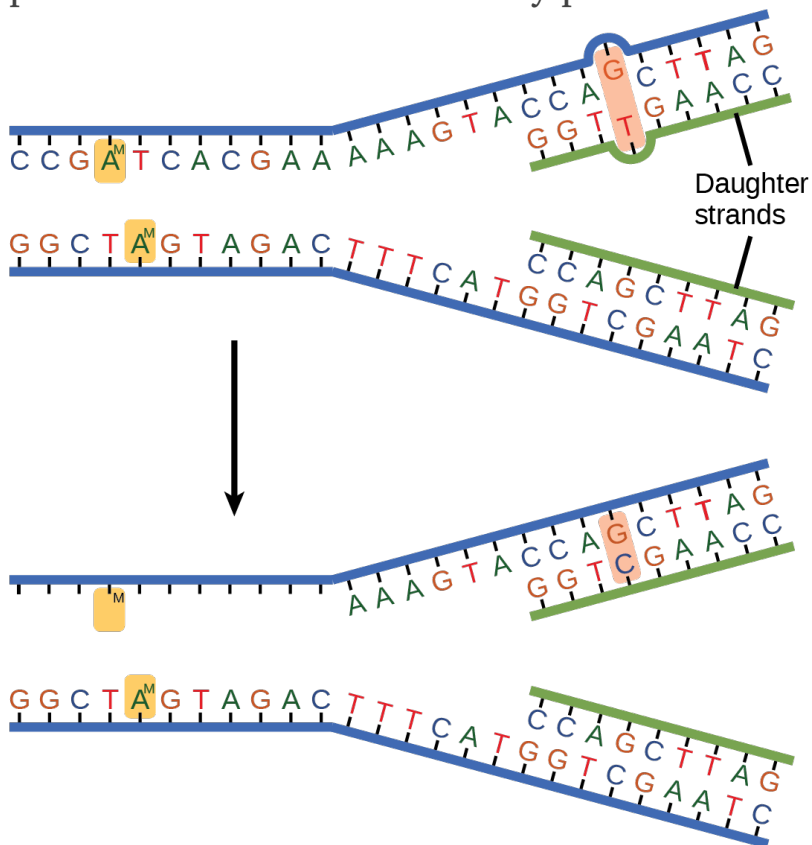
Proofreading by DNA polymerase corrects errors during replication.



Some errors are not corrected during replication, but are instead corrected after replication is completed; this type of repair is known as **mismatch repair** ([\[link\]](#)). Specific repair enzymes recognize the mismatched nucleotide and excise part of the strand that contains it; the excised region is then resynthesized. If the mismatch remains uncorrected, it may lead to more permanent damage when the mismatched DNA is replicated. How do mismatch repair enzymes recognize which of the two bases is the incorrect one? In *E. coli*, after replication, the nitrogenous base adenine acquires a methyl group; the parental DNA strand will have methyl groups, whereas the newly synthesized strand lacks them. Thus, DNA polymerase is able to remove the wrongly incorporated bases from the newly synthesized, non-methylated strand. In eukaryotes, the mechanism is not very well understood, but it is believed to involve recognition of unsealed nicks in the new strand, as well as a short-term continuing association of some of the replication proteins with the new daughter strand

after replication has completed.

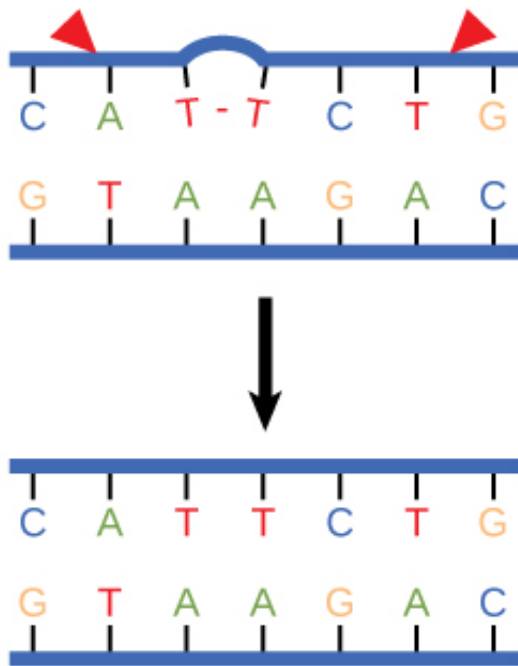
In mismatch repair, the incorrectly added base is detected after replication. The mismatch repair proteins detect this base and remove it from the newly synthesized strand by nuclease action. The gap is now filled with the correctly paired base.



Another type of repair mechanism, **nucleotide excision repair**, is similar to mismatch repair, except that it is used to remove damaged bases rather than mismatched ones. The repair enzymes replace abnormal bases by making a cut on both the 3' and 5' ends of the damaged base ([\[link\]](#)). The

segment of DNA is removed and replaced with the correctly paired nucleotides by the action of DNA pol. Once the bases are filled in, the remaining gap is sealed with a phosphodiester linkage catalyzed by DNA ligase. This repair mechanism is often employed when UV exposure causes the formation of pyrimidine dimers.

Nucleotide excision repairs thymine dimers. When exposed to UV light, thymines lying adjacent to each other can form thymine dimers. In normal cells, they are excised and replaced.



A well-studied example of mistakes not being corrected is seen in people suffering from xeroderma pigmentosa ([link](#)). Affected individuals have skin that is highly sensitive to UV rays from the sun.

When individuals are exposed to UV light, pyrimidine dimers, especially those of thymine, are formed; people with xeroderma pigmentosa are not able to repair the damage. These are not repaired because of a defect in the nucleotide excision repair enzymes, whereas in normal individuals, the thymine dimers are excised and the defect is corrected. The thymine dimers distort the structure of the DNA double helix, and this may cause problems during DNA replication. People with xeroderma pigmentosa may have a higher risk of contracting skin cancer than those who don't have the condition.

Xeroderma pigmentosa is a condition in which thymine dimerization from exposure to UV light is not repaired. Exposure to sunlight results in skin lesions. (credit: James Halpern et al.)



Errors during DNA replication are not the only reason why mutations arise in DNA. **Mutations**, variations in the nucleotide sequence of a genome, can also occur because of damage to DNA. Such mutations may be of two types: induced or spontaneous. **Induced mutations** are those that result from an exposure to chemicals, UV rays, x-rays, or some other environmental agent. **Spontaneous mutations** occur without any exposure to any environmental agent; they are a result of natural reactions taking place within the body.

Mutations may have a wide range of effects. Point mutations are those mutations that affect a single

base pair. The most common nucleotide mutations are substitutions, in which one base is replaced by another. These substitutions can be of two types, either transitions or transversions. **Transition substitution** refers to a purine or pyrimidine being replaced by a base of the same kind; for example, a purine such as adenine may be replaced by the purine guanine. **Transversion substitution** refers to a purine being replaced by a pyrimidine, or vice versa; for example, cytosine, a pyrimidine, is replaced by adenine, a purine. Some point mutations are not expressed; these are known as silent mutations. Silent mutations are usually due to a substitution in the third base of a codon, which often represents the same amino acid as the original codon. Other point mutations can result in the replacement of one amino acid by another, which may alter the function of the protein. Point mutations that generate a stop codon can terminate a protein early.

Some mutations can result in an increased number of copies of the same codon. These are called trinucleotide repeat expansions and result in repeated regions of the same amino acid. Mutations can also be the result of the addition of a base, known as an insertion, or the removal of a base, also known as deletion. If an insertion or deletion results in the alteration of the translational reading frame (a frameshift mutation), the resultant protein is usually nonfunctional. Sometimes a piece of DNA

from one chromosome may get translocated to another chromosome or to another region of the same chromosome; this is also known as translocation. These mutation types are shown in [\[link\]](#).

Visual Connection

Mutations can lead to changes in the protein sequence encoded by the DNA.

Point Mutations

Silent: has no effect on the protein sequence



Missense: results in an amino acid substitution

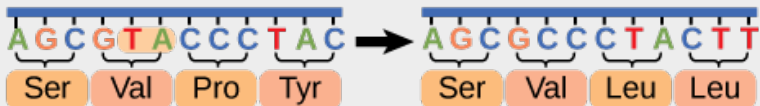


Nonsense: substitutes a stop codon for an amino acid



Frameshift Mutations

Insertions or deletions of nucleotides may result in a shift in the reading frame or insertion of a stop codon.



A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

Mutations in repair genes have been known to cause cancer. Many mutated repair genes have been

implicated in certain forms of pancreatic cancer, colon cancer, and colorectal cancer. Mutations can affect either somatic cells or germ cells. If many mutations accumulate in a somatic cell, they may lead to problems such as the uncontrolled cell division observed in cancer. If a mutation takes place in germ cells, the mutation will be passed on to the next generation, as in the case of hemophilia and xeroderma pigmentosa.

Section Summary

DNA polymerase can make mistakes while adding nucleotides. It edits the DNA by proofreading every newly added base. Incorrect bases are removed and replaced by the correct base before proceeding with elongation. Most mistakes are corrected during replication, although when this does not happen, the mismatch repair mechanism is employed. Mismatch repair enzymes recognize the wrongly incorporated base and excise it from the DNA, replacing it with the correct base. In yet another type of repair, nucleotide excision repair, a damaged base is removed along with a few bases on the 5' and 3' end, and these are replaced by copying the template with the help of DNA polymerase. The ends of the newly synthesized fragment are attached to the rest of the DNA using DNA ligase, which creates a phosphodiester bond.

Most mistakes are corrected, and if they are not, they may result in a mutation, defined as a permanent change in the DNA sequence. Mutations can be of many types, such as substitution, deletion, insertion, and trinucleotide repeat expansions. Mutations in repair genes may lead to serious consequences such as cancer. Mutations can be induced or may occur spontaneously.

Visual Connection Questions

[\[link\]](#) A frameshift mutation that results in the insertion of three nucleotides is often less deleterious than a mutation that results in the insertion of one nucleotide. Why?

[\[link\]](#) If three nucleotides are added, one additional amino acid will be incorporated into the protein chain, but the reading frame won't shift.

Review Questions

During proofreading, which of the following

enzymes reads the DNA?

1. primase
 2. topoisomerase
 3. DNA pol
 4. helicase
-

C

The initial mechanism for repairing nucleotide errors in DNA is _____.

1. mismatch repair
 2. DNA polymerase proofreading
 3. nucleotide excision repair
 4. thymine dimers
-

B

A scientist creates fruit fly larvae with a mutation that eliminates the exonuclease function of DNA pol III. Which prediction about the mutational load in the adult fruit flies is most likely to be correct?

1. The adults with the DNA pol III mutation will have significantly more mutations than average.

2. The adults with the DNA pol III mutation will have slightly more mutations than average.
 3. The adults with the DNA pol III mutation will have the same number of mutations as average.
 4. The adults with the DNA pol III mutation will have fewer mutations than average.
-

B

Critical Thinking Questions

What is the consequence of mutation of a mismatch repair enzyme? How will this affect the function of a gene?

Mutations are not repaired, as in the case of xeroderma pigmentosa. Gene function may be affected or it may not be expressed.

An adult with a history of tanning has his genome sequenced. The beginning of a protein-coding region of his DNA reads ATGGGGATATGGCAT. If the protein-coding

region of a healthy adult reads
ATGGGGATATGAGCAT, identify the site and
type of mutation.

This is a frameshift mutation with a deletion of
an “A” in the 12th position of the coding region.

Patient: ATGGGGATATGGCAT

Normal: ATGGGGATATGAGCAT

Glossary

induced mutation

mutation that results from exposure to
chemicals or environmental agents

mutation

variation in the nucleotide sequence of a
genome

mismatch repair

type of repair mechanism in which
mismatched bases are removed after
replication

nucleotide excision repair

type of DNA repair mechanism in which the
wrong base, along with a few nucleotides
upstream or downstream, are removed

proofreading

function of DNA pol in which it reads the newly added base before adding the next one

point mutation

mutation that affects a single base

silent mutation

mutation that is not expressed

spontaneous mutation

mutation that takes place in the cells as a result of chemical reactions taking place naturally without exposure to any external agent

transition substitution

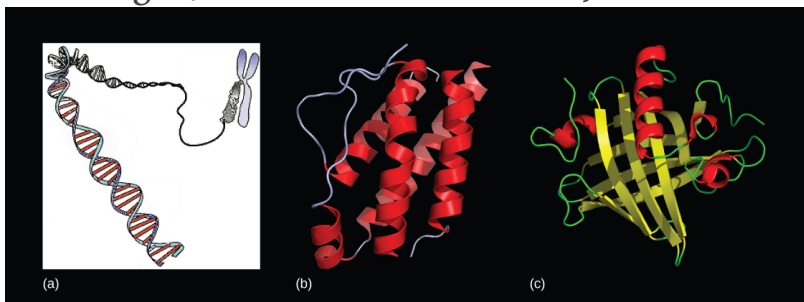
when a purine is replaced with a purine or a pyrimidine is replaced with another pyrimidine

transversion substitution

when a purine is replaced by a pyrimidine or a pyrimidine is replaced by a purine

Introduction

class = "introduction" Genes, which are carried on (a) chromosomes, are linearly organized instructions for making the RNA and protein molecules that are necessary for all of the processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes. (credit “chromosome: National Human Genome Research Institute; credit “interleukin-2”: Ramin Herati/Created from PDB 1M47 and rendered with Pymol; credit “alpha-2u-globulin”: Darren Logan/rendered with AISMIG)



Since the rediscovery of Mendel’s work in 1900, the definition of the gene has progressed from an abstract unit of heredity to a tangible molecular entity capable of replication, expression, and mutation ([\[link\]](#)). Genes are composed of DNA and are linearly arranged on chromosomes. Genes specify the sequences of amino acids, which are the building blocks of proteins. In turn, proteins are responsible for orchestrating nearly every function of the cell. Both genes and the proteins they encode

are absolutely essential to life as we know it.




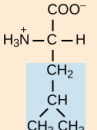
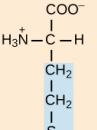
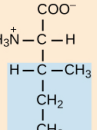



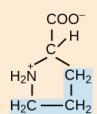
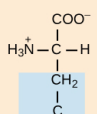
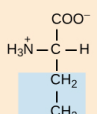
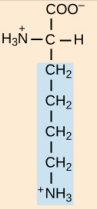
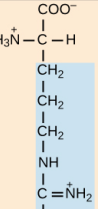
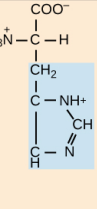
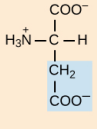
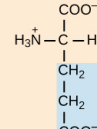
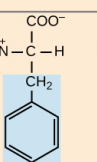
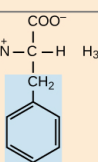
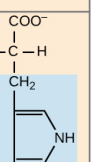
The Genetic Code

By the end of this section, you will be able to do the following:

- Explain the “central dogma” of DNA-protein synthesis
- Describe the genetic code and how the nucleotide sequence prescribes the amino acid and the protein sequence

The cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template on ribosomes converts nucleotide-based genetic information into a protein product. That is the central dogma of DNA-protein synthesis. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 “letters” ([\[link\]](#)). Different amino acids have different chemistries (such as acidic versus basic, or polar and nonpolar) and different structural constraints. Variation in amino acid sequence is responsible for the enormous variation in protein structure and function. Structures of the 20 amino acids found in proteins are shown. Each amino acid is composed of an amino group (NH_3^+), a carboxyl group (COO^-), and a side chain (blue). The side chain may be nonpolar, polar, or charged, as well as large or small. It is the variety of amino acid side chains that

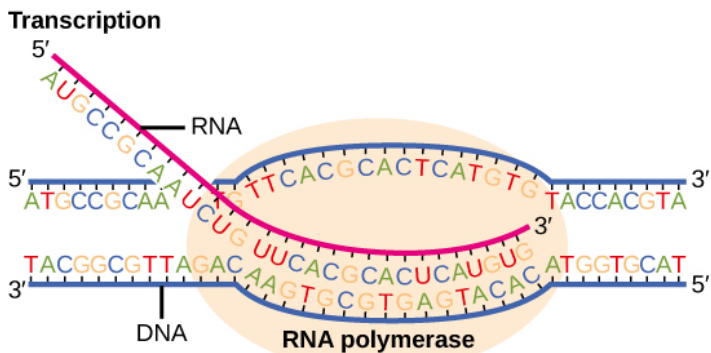
gives rise to the incredible variation of protein structure and function.

AMINO ACID			
Nonpolar, aliphatic R groups			
	Glycine	Alanine	Valine
			
	Leucine	Methionine	Isoleucine
			
	Serine	Threonine	Cysteine
Polar, uncharged R groups			
	Proline	Asparagine	Glutamine
AMINO ACID			
Positively charged R groups			
	Lysine	Arginine	Histidine
Negatively charged R groups			
	Aspartate	Glutamate	
Nonpolar, aromatic R groups			
	Phenylalanine	Tyrosine	Tryptophan

Instructions on DNA are transcribed onto messenger RNA. Ribosomes are able to read the genetic information inscribed on a strand of messenger RNA and use this information to string amino acids together into a protein. This figure shows the genetic code for translating each nucleotide triplet in mRNA into an amino acid or a termination signal in a protein. (credit: modification of work by NIH)The deletion of two nucleotides shifts the reading frame of an mRNA and changes the entire protein message, creating a nonfunctional protein or terminating protein synthesis altogether.

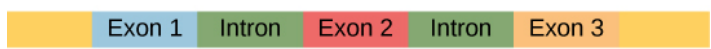
The Central Dogma: DNA Encodes RNA; RNA Encodes Protein

The flow of genetic information in cells from DNA to mRNA to protein is described by the **central dogma** ([\[link\]](#)), which states that genes specify the sequence of mRNAs, which in turn specify the sequence of amino acids making up all proteins. The decoding of one molecule to another is performed by specific proteins and RNAs. Because the information stored in DNA is so central to cellular function, it makes intuitive sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected. The copying of DNA to RNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every nucleotide read in the DNA strand. The translation to protein is a bit more complex because three mRNA nucleotides correspond to one amino acid in the polypeptide sequence. However, the translation to protein is still systematic and **colinear**, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

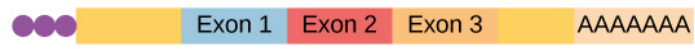


RNA processing

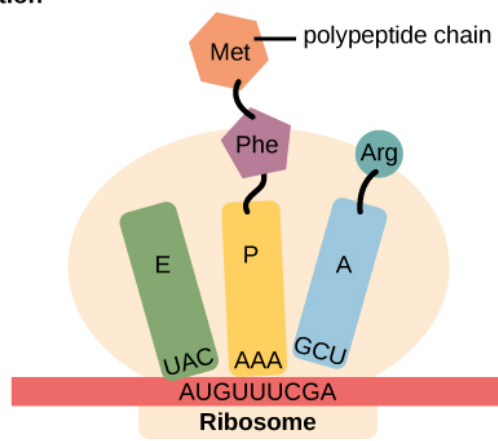
Primary RNA transcript



Spliced RNA



Translation



The Genetic Code Is Degenerate and Universal

Each amino acid is defined by a three-nucleotide sequence called the triplet codon. Given the different numbers of “letters” in the mRNA and protein “alphabets,” scientists theorized that single amino acids must be represented by combinations of nucleotides. Nucleotide doublets would not be sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations (4²). In contrast, there are 64 possible nucleotide triplets (4³), which is far more than the number of amino acids. Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was “**degenerate**.” In other words, a given amino acid could be encoded by more than one nucleotide triplet. This was later confirmed experimentally: Francis Crick and Sydney Brenner used the chemical mutagen proflavin to insert one, two, or three nucleotides into the gene of a virus. When one or two nucleotides were inserted, the normal proteins were not produced. When three nucleotides were inserted, the protein was synthesized and functional. This demonstrated that the amino acids must be specified by groups of three nucleotides. These nucleotide triplets are called **codons**. The insertion of one or two nucleotides completely changed the triplet **reading frame**, thereby altering the message for every subsequent amino acid ([\[link\]](#)). Though insertion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

Scientists painstakingly solved the genetic code by translating synthetic mRNAs in vitro and sequencing the proteins they specified ([\[link\]](#)).

		Second letter					
		U	C	A	G		
First letter	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA Stop UAG Stop	UGU } Cys UGC } UGA Stop UGG Trp	U C A G	Third letter
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }	U C A G	
	A	AUU } AUC } Ile AUA } AUG Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G	
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }	U C A G	

In addition to codons that instruct the addition of a specific amino acid to a polypeptide chain, three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **nonsense codons**, or *stop codons*. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the start codon to initiate translation. The **reading frame** for translation is set by the AUG start codon near the 5' end of the mRNA. Following the start codon, the mRNA is read in groups of three until a stop codon

is encountered.

The arrangement of the coding table reveals the structure of the code. There are sixteen "blocks" of codons, each specified by the first and second nucleotides of the codons within the block, e.g., the "AC*" block that corresponds to the amino acid threonine (Thr). Some blocks are divided into a pyrimidine half, in which the codon ends with U or C, and a purine half, in which the codon ends with A or G. Some amino acids get a whole block of four codons, like alanine (Ala), threonine (Thr) and proline (Pro). Some get the pyrimidine half of their block, like histidine (His) and asparagine (Asn). Others get the purine half of their block, like glutamate (Glu) and lysine (Lys). Note that some amino acids get a block and a half-block for a total of six codons.

The specification of a single amino acid by multiple similar codons is called "degeneracy." Degeneracy is believed to be a cellular mechanism to reduce the negative impact of random mutations. Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. For example, aspartate (Asp) and glutamate (Glu), which occupy the GA* block, are both negatively charged. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might specify the same amino acid but

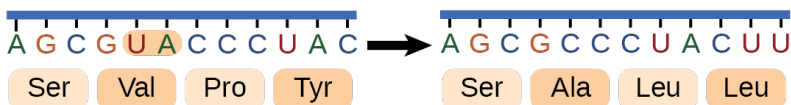
have no effect or specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

The genetic code is nearly universal. With a few minor exceptions, virtually all species use the same genetic code for protein synthesis. Conservation of codons means that a purified mRNA encoding the globin protein in horses could be transferred to a tulip cell, and the tulip would synthesize horse globin. That there is only one genetic code is powerful evidence that all of life on Earth shares a common origin, especially considering that there are about 10^4 possible combinations of 20 amino acids and 64 triplet codons.

Link to Learning

Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this [site](#).

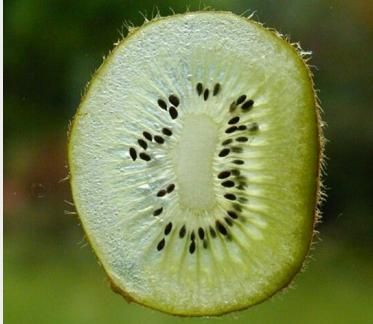
Frameshift Mutations



Scientific Method Connection

Which Has More DNA: A Kiwi or a Strawberry?

Do you think that a kiwi or a strawberry has more DNA per fruit? (credit “kiwi”: "Kelbv"/Flickr; credit: “strawberry”: Alisdair McDiarmid)



Question: Would a kiwi and strawberry that are approximately the same size ([\[link\]](#)) also have approximately the same amount of DNA?

Background: Genes are carried on chromosomes and are made of DNA. All mammals are diploid, meaning they have two copies of each chromosome. However, not all plants are diploid. The common strawberry is octoploid ($8n$) and the cultivated kiwi is hexaploid ($6n$). Research the total number of chromosomes in the cells of each of these fruits and think about how this might correspond to the amount of DNA in these fruits' cell nuclei. What other factors might contribute to the total amount of DNA in a single fruit? Read about the technique of DNA isolation to understand how each step in the isolation protocol helps liberate and precipitate DNA.

Hypothesis: Hypothesize whether you would be able to detect a difference in DNA quantity from

similarly sized strawberries and kiwis. Which fruit do you think would yield more DNA?

Test your hypothesis: Isolate the DNA from a strawberry and a kiwi that are similarly sized. Perform the experiment in at least triplicate for each fruit

1. Prepare a bottle of DNA extraction buffer from 900 mL water, 50 mL dish detergent, and two teaspoons of table salt. Mix by inversion (cap it and turn it upside down a few times).
2. Grind a strawberry and a kiwi by hand in a plastic bag, or using a mortar and pestle, or with a metal bowl and the end of a blunt instrument. Grind for at least two minutes per fruit.
3. Add 10 mL of the DNA extraction buffer to each fruit, and mix well for at least one minute.
4. Remove cellular debris by filtering each fruit mixture through cheesecloth or porous cloth and into a funnel placed in a test tube or an appropriate container.
5. Pour ice-cold ethanol or isopropanol (rubbing alcohol) into the test tube. You should observe white, precipitated DNA.
6. Gather the DNA from each fruit by winding it around separate glass rods.

Record your observations: Because you are not quantitatively measuring DNA volume, you can

record for each trial whether the two fruits produced the same or different amounts of DNA as observed by eye. If one or the other fruit produced noticeably more DNA, record this as well.

Determine whether your observations are consistent with several pieces of each fruit.

Analyze your data: Did you notice an obvious difference in the amount of DNA produced by each fruit? Were your results reproducible?

Draw a conclusion: Given what you know about the number of chromosomes in each fruit, can you conclude that chromosome number necessarily correlates to DNA amount? Can you identify any drawbacks to this procedure? If you had access to a laboratory, how could you standardize your comparison and make it more quantitative?

Section Summary

The genetic code refers to the DNA alphabet (A, T, C, G), the RNA alphabet (A, U, C, G), and the polypeptide alphabet (20 amino acids). The central dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is degenerate

because 64 triplet codons in mRNA specify only 20 amino acids and three nonsense codons. Most amino acids have several similar codons. Almost every species on the planet uses the same genetic code.

Review Questions

The AUC and AUA codons in mRNA both specify isoleucine. What feature of the genetic code explains this?

1. complementarity
2. nonsense codons
3. universality
4. degeneracy

D

How many nucleotides are in 12 mRNA codons?

1. 12
 2. 24
 3. 36
 4. 48
-

C

Which event contradicts the central dogma of molecular biology?

1. Poly-A polymerase enzymes process mRNA in the nucleus.
2. Endonuclease enzymes splice out and repair damaged DNA.
3. Scientists use reverse transcriptase enzymes to make DNA from RNA.
4. Codons specifying amino acids are degenerate and universal.

C

Critical Thinking Questions

Imagine if there were 200 commonly occurring amino acids instead of 20. Given what you know about the genetic code, what would be the shortest possible codon length? Explain.

For 200 commonly occurring amino acids, codons consisting of four types of nucleotides

would have to be at least four nucleotides long, because $4^4 = 256$. There would be much less degeneracy in this case.

Discuss how degeneracy of the genetic code makes cells more robust to mutations.

Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid and have no effect, or may specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

A scientist sequencing mRNA identifies the following strand:

CUAUGUGUCGUAACAGCCGAUGACCCG

What is the sequence of the amino acid chain this mRNA makes when it is translated?

Met Cys Arg Asn Ser Arg

The first step to writing the amino acid

sequence is to find the start codon AUG. Then, the nucleotide sequence is separated into triplets: CU AUG UGU CGU AAC AGC CGA UGA. We stop the translation at UGA because that triplet encodes a stop codon. When we convert these codons to amino acids, the sequence becomes Met Cys Arg Asn Ser Arg.

Glossary

central dogma

states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins

codon

three consecutive nucleotides in mRNA that specify the insertion of an amino acid or the release of a polypeptide chain during translation

colinear

in terms of RNA and protein, three “units” of RNA (nucleotides) specify one “unit” of protein (amino acid) in a consecutive fashion

degeneracy

(of the genetic code) describes that a given amino acid can be encoded by more than one nucleotide triplet; the code is degenerate, but not ambiguous

nonsense codon

one of the three mRNA codons that specifies termination of translation

reading frame

sequence of triplet codons in mRNA that specify a particular protein; a ribosome shift of one or two nucleotides in either direction completely abolishes synthesis of that protein

Prokaryotic Transcription

By the end of this section, you will be able to do the following:

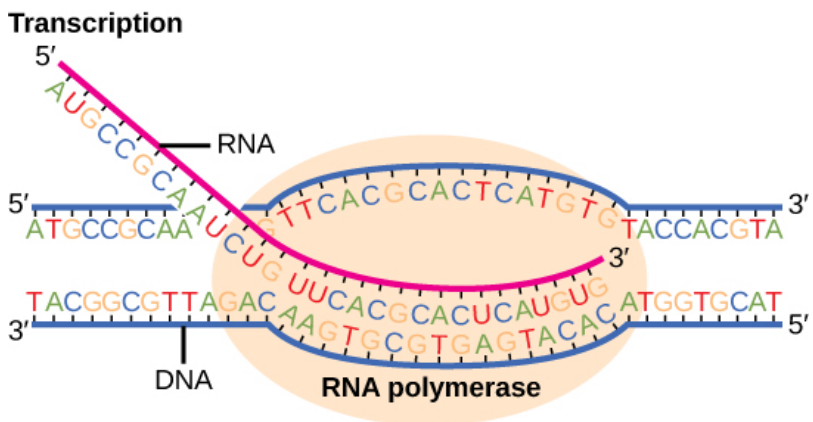
- List the different steps in prokaryotic transcription
- Discuss the role of promoters in prokaryotic transcription
- Describe how and when transcription is terminated

The prokaryotes, which include Bacteria and Archaea, are mostly single-celled organisms that, by definition, lack membrane-bound nuclei and other organelles. A bacterial chromosome is a closed circle that, unlike eukaryotic chromosomes, is not organized around histone proteins. The central region of the cell in which prokaryotic DNA resides is called the nucleoid region. In addition, prokaryotes often have abundant **plasmids**, which are shorter, circular DNA molecules that may only contain one or a few genes. Plasmids can be transferred independently of the bacterial chromosome during cell division and often carry traits such as those involved with antibiotic resistance.

Transcription in prokaryotes (and in eukaryotes) requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a **transcription bubble**.

Transcription always proceeds from the same DNA strand for each gene, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **nontemplate strand**, or the coding strand. The only nucleotide difference is that in mRNA, all of the T nucleotides are replaced with U nucleotides ([\[link\]](#)). In an RNA double helix, A can bind U via two hydrogen bonds, just as in A–T pairing in a DNA double helix.

Messenger RNA is a copy of protein-coding information in the coding strand of DNA, with the substitution of U in the RNA for T in the coding sequence. However, new RNA nucleotides base pair with the nucleotides of the template strand. RNA is synthesized in its 5'-3' direction, using the enzyme RNA polymerase. As the template is read, the DNA unwinds ahead of the polymerase and then rewinds behind it.



The nucleotide pair in the DNA double helix that corresponds to the site from which the first 5' mRNA

nucleotide is transcribed is called the +1 site, or the **initiation site**. Nucleotides preceding the initiation site are denoted with a “-” and are designated *upstream nucleotides*. Conversely, nucleotides following the initiation site are denoted with “+” numbering and are called *downstream nucleotides*. The σ subunit of prokaryotic RNA polymerase recognizes consensus sequences found in the promoter region upstream of the transcription start site. The σ subunit dissociates from the polymerase after transcription has been initiated.

Initiation of Transcription in Prokaryotes

Prokaryotes do not have membrane-enclosed nuclei. Therefore, the processes of transcription, translation, and mRNA degradation can all occur simultaneously. The intracellular level of a bacterial protein can quickly be amplified by multiple transcription and translation events that occur concurrently on the same DNA template.

Prokaryotic genomes are very compact, and prokaryotic transcripts often cover more than one gene or cistron (a coding sequence for a single protein). Polycistronic mRNAs are then translated to produce more than one kind of protein.

Our discussion here will exemplify transcription by describing this process in *Escherichia coli*, a well-studied eubacterial species. Although some differences exist between transcription in *E. coli* and

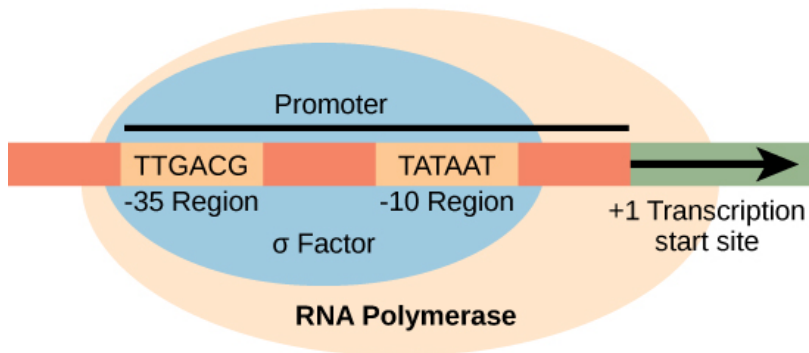
transcription in archaea, an understanding of *E. coli* transcription can be applied to virtually all bacterial species.

Prokaryotic RNA Polymerase

Prokaryotes use the same RNA polymerase to transcribe all of their genes. In *E. coli*, the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α , α , β , and β' , comprise the polymerase **core enzyme**. These subunits assemble every time a gene is transcribed, and they disassemble once transcription is complete. Each subunit has a unique role; the two α -subunits are necessary to assemble the polymerase on the DNA; the β -subunit binds to the ribonucleoside triphosphate that will become part of the nascent mRNA molecule; and the β' subunit binds the DNA template strand. The fifth subunit, σ , is involved only in transcription initiation. It confers transcriptional specificity such that the polymerase begins to synthesize mRNA from an appropriate initiation site. Without σ , the core enzyme would transcribe from random sites and would produce mRNA molecules that specified protein gibberish. The polymerase comprised of all five subunits is called the **holoenzyme**.

Prokaryotic Promoters

A **promoter** is a DNA sequence onto which the transcription machinery, including RNA polymerase, binds and initiates transcription. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all the time, some of the time, or infrequently. Although promoters vary among prokaryotic genomes, a few elements are evolutionarily conserved in many species. At the -10 and -35 regions upstream of the initiation site, there are two *promoter consensus sequences*, or regions that are similar across all promoters and across various bacterial species ([\[link\]](#)). The -10 sequence, called the -10 region, has the consensus sequence TATAAT. The -35 sequence has the consensus sequence TTGACA. These consensus sequences are recognized and bound by σ . Once this interaction is made, the subunits of the core enzyme bind to the site. The A-T-rich -10 region facilitates unwinding of the DNA template, and several phosphodiester bonds are made. The transcription initiation phase ends with the production of abortive transcripts, which are polymers of approximately 10 nucleotides that are made and released.



Link to Learning

View this [MolecularMovies animation](#) to see the first part of transcription and the base sequence repetition of the TATA box.

Elongation and Termination in Prokaryotes

The transcription elongation phase begins with the release of the σ subunit from the polymerase. The dissociation of σ allows the core enzyme to proceed along the DNA template, synthesizing mRNA in the 5' to 3' direction at a rate of approximately 40 nucleotides per second. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it. The base pairing

between DNA and RNA is not stable enough to maintain the stability of the mRNA synthesis components. Instead, the RNA polymerase acts as a stable linker between the DNA template and the nascent RNA strands to ensure that elongation is not interrupted prematurely.

Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

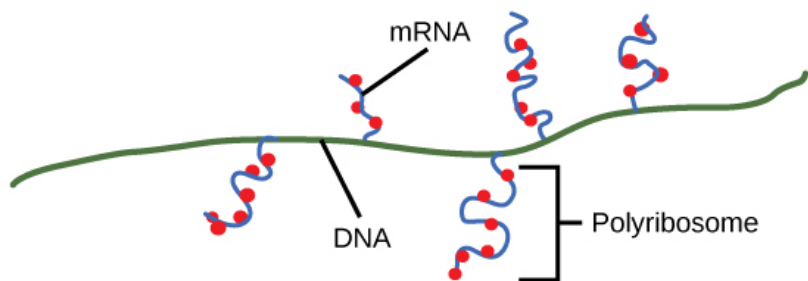
Prokaryotic Termination Signals

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals. One is protein-based and the other is RNA-based. **Rho-dependent termination** is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase encounters a run of G nucleotides on the DNA template and it stalls. As a result, the rho protein collides with the polymerase. The interaction with rho releases the mRNA from the transcription bubble.

Rho-independent termination is controlled by

specific sequences in the DNA template strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. The mRNA folds back on itself, and the complementary C–G nucleotides bind together. The result is a **stable hairpin** that causes the polymerase to stall as soon as it begins to transcribe a region rich in A–T nucleotides. The complementary U–A region of the mRNA transcript forms only a weak interaction with the template DNA. This, coupled with the stalled polymerase, induces enough instability for the core enzyme to break away and liberate the new mRNA transcript.

Upon termination, the process of transcription is complete. By the time termination occurs, the prokaryotic transcript would already have been used to begin synthesis of numerous copies of the encoded protein because these processes can occur concurrently. The unification of transcription, translation, and even mRNA degradation is possible because all of these processes occur in the same 5' to 3' direction, and because there is no membranous compartmentalization in the prokaryotic cell ([\[link\]](#)). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.



Link to Learning

Visit this [BioStudio animation](#) to see the process of prokaryotic transcription.

Section Summary

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template comprising two consensus sequences that recruit RNA polymerase. The prokaryotic polymerase consists of a core enzyme of four protein subunits and a σ protein that assists only with initiation. Elongation synthesizes mRNA in the 5' to 3' direction at a rate of 40 nucleotides per second. Termination liberates the mRNA and occurs either by rho protein interaction or by the formation of an mRNA hairpin.

Review Questions

Which subunit of the *E. coli* polymerase confers specificity to transcription?

1. α
2. β
3. β'
4. σ

D

The -10 and -35 regions of prokaryotic promoters are called consensus sequences because _____.

1. they are identical in all bacterial species
2. they are similar in all bacterial species
3. they exist in all organisms
4. they have the same function in all organisms

B

Three different bacteria species have the following consensus sequences upstream of a conserved gene.

--	--	--	--	--

	Species A	Species B	Species C
10	TAATAAT	TTTAAT	TATATT
-35	TTGACA	TTGGCC	TTGAAA

Order the bacteria from most to least efficient initiation of gene transcription.

1. $A > B > C$
2. $B > C > A$
3. $C > B > A$
4. $A > C > B$

D

Critical Thinking Questions

If mRNA is complementary to the DNA template strand and the DNA template strand is complementary to the DNA nontemplate strand, then why are base sequences of mRNA and the DNA nontemplate strand not identical? Could they ever be?

DNA is different from RNA in that T nucleotides in DNA are replaced with U nucleotides in RNA. Therefore, they could never be identical in base

sequence.

In your own words, describe the difference between rho-dependent and rho-independent termination of transcription in prokaryotes.

Rho-dependent termination is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase stalls at a run of G nucleotides on the DNA template. The rho protein collides with the polymerase and releases mRNA from the transcription bubble. Rho-independent termination is controlled by specific sequences in the DNA template strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. This creates an mRNA hairpin that causes the polymerase to stall right as it begins to transcribe a region rich in A–T nucleotides. Because A–U bonds are less thermostable, the core enzyme falls away.

A fragment of bacterial DNA reads:

3' –

TACCTATAATCTCAATTGATAGAAGCACTCTAC–

5'

Assuming that this fragment is the template strand, what is the sequence of mRNA that would be transcribed? (Hint: Be sure to identify the initiation site.)

ACUAUCUUCGUGAGAUG

By examining the DNA sequence, we can see that there is a -10 consensus sequence near the 3' end of the fragment. If we then count downstream, the +1 initiation site is the T immediately following the sequence AAT. This means the DNA fragment that will serve as the template for transcription has the sequence TGATAGAAGCACTCTAC. The mRNA made from this template will have complimentary base pairing with uracil (U) instead of thymine (T). This gives us ACUAUCUUCGUGAGAUG as the transcribed mRNA sequence.

Glossary

consensus

DNA sequence that is used by many species to perform the same or similar functions

core enzyme

prokaryotic RNA polymerase consisting of α , α , β , and β' but missing σ ; this complex performs elongation

downstream

nucleotides following the initiation site in the direction of mRNA transcription; in general, sequences that are toward the 3' end relative to a site on the mRNA

hairpin

structure of RNA when it folds back on itself and forms intramolecular hydrogen bonds between complementary nucleotides

holoenzyme

prokaryotic RNA polymerase consisting of α , α , β , β' , and σ ; this complex is responsible for transcription initiation

initiation site

nucleotide from which mRNA synthesis proceeds in the 5' to 3' direction; denoted with a “+ 1”

nontemplate strand

strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

plasmid

extrachromosomal, covalently closed, circular DNA molecule that may only contain one or a few genes; common in prokaryotes

promoter

DNA sequence to which RNA polymerase and associated factors bind and initiate transcription

rho-dependent termination

in prokaryotes, termination of transcription by an interaction between RNA polymerase and the rho protein at a run of G nucleotides on the DNA template

rho-independent

termination sequence-dependent termination of prokaryotic mRNA synthesis; caused by hairpin formation in the mRNA that stalls the polymerase

TATA box

conserved promoter sequence in eukaryotes and prokaryotes that helps to establish the initiation site for transcription

template strand

strand of DNA that specifies the complementary mRNA molecule

transcription bubble

region of locally unwound DNA that allows for transcription of mRNA

upstream

nucleotides preceding the initiation site; in

general, sequences toward the 5' end relative to a site on the mRNA

Eukaryotic Transcription

By the end of this section, you will be able to do the following:

- List the steps in eukaryotic transcription
- Discuss the role of RNA polymerases in transcription
- Compare and contrast the three RNA polymerases
- Explain the significance of transcription factors

Prokaryotes and eukaryotes perform fundamentally the same process of transcription, with a few key differences. The most important difference between prokaryote and eukaryote transcription is due to the latter's membrane-bound nucleus and organelles. With the genes bound in a nucleus, the eukaryotic cell must be able to transport its mRNA to the cytoplasm and must protect its mRNA from degrading before it is translated. Eukaryotes also employ three different polymerases that each transcribe a different subset of genes. Eukaryotic mRNAs are usually *monogenic*, meaning that they specify a single protein.

Initiation of Transcription in Eukaryotes

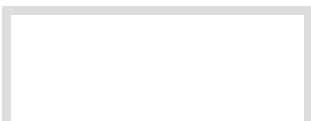
Unlike the prokaryotic polymerase that can bind to a DNA template on its own, eukaryotes require

several other proteins, called transcription factors, to first bind to the promoter region and then to help recruit the appropriate polymerase.

The Three Eukaryotic RNA Polymerases

The features of eukaryotic mRNA synthesis are markedly more complex than those of prokaryotes. Instead of a single polymerase comprising five subunits, the eukaryotes have three polymerases that are each made up of 10 subunits or more. Each eukaryotic polymerase also requires a distinct set of transcription factors to bring it to the DNA template.

RNA polymerase I is located in the nucleolus, a specialized nuclear substructure in which ribosomal RNA (rRNA) is transcribed, processed, and assembled into ribosomes ([\[link\]](#)). The rRNA molecules are considered structural RNAs because they have a cellular role but are not translated into protein. The rRNAs are components of the ribosome and are essential to the process of translation. RNA polymerase I synthesizes all of the rRNAs from the tandemly duplicated set of 18S, 5.8S, and 28S ribosomal genes. (Note that the “S” designation applies to “Svedberg” units, a nonadditive value that characterizes the speed at which a particle sediments during centrifugation.)



Locations, Products, and Sensitivities of the Three Eukaryotic RNA Polymerases

RNA Polymerase	Cellular Compartment	Product of Transcription	α -Amanitin Sensitivity
I	Nucleolus	All rRNAs except 5S rRNA	Insensitive
II	Nucleus	All protein-coding nuclear pre-mRNAs	Extremely sensitive
III	Nucleus	5S rRNA, tRNAs, and small nuclear RNAs	Moderately sensitive

RNA polymerase II is located in the nucleus and synthesizes all protein-coding nuclear pre-mRNAs. Eukaryotic pre-mRNAs undergo extensive processing after transcription but before translation. For clarity, this module's discussion of transcription and translation in eukaryotes will use the term "mRNAs" to describe only the mature, processed molecules that are ready to be translated. RNA polymerase II is

responsible for transcribing the overwhelming majority of eukaryotic genes.

RNA polymerase III is also located in the nucleus. This polymerase transcribes a variety of structural RNAs that includes the 5S pre-rRNA, transfer pre-RNAs (pre-tRNAs), and **small nuclear pre-RNAs**. The tRNAs have a critical role in translation; they serve as the “adaptor molecules” between the mRNA template and the growing polypeptide chain. Small nuclear RNAs have a variety of functions, including “splicing” pre-mRNAs and regulating transcription factors.

A scientist characterizing a new gene can determine which polymerase transcribes it by testing whether the gene is expressed in the presence of α -amanitin, an oligopeptide toxin produced by the fly agaric toadstool mushroom and other species of *Amanita*. Interestingly, the α -amanitin affects the three polymerases very differently ([\[link\]](#)). RNA polymerase I is completely insensitive to α -amanitin, meaning that the polymerase can transcribe DNA in vitro in the presence of this poison. RNA polymerase III is moderately sensitive to the toxin. In contrast, RNA polymerase II is extremely sensitive to α -amanitin. The toxin prevents the enzyme from progressing down the DNA, and thus inhibits transcription. Knowing the transcribing polymerase can provide clues as to the general function of the gene being studied. Because

RNA polymerase II transcribes the vast majority of genes, we will focus on this polymerase in our subsequent discussions about eukaryotic transcription factors and promoters.

RNA Polymerase II Promoters and Transcription Factors

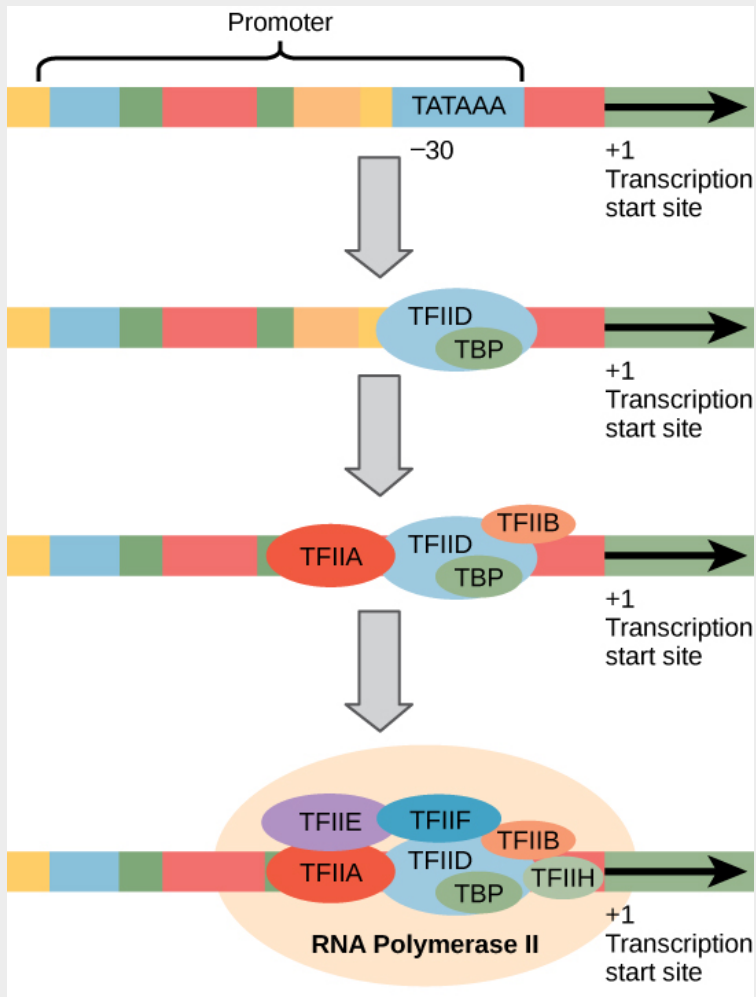
Eukaryotic promoters are much larger and more intricate than prokaryotic promoters. However, both have a sequence similar to the -10 sequence of prokaryotes. In eukaryotes, this sequence is called the TATA box, and has the consensus sequence TATAAA on the coding strand. It is located at -25 to -35 bases relative to the initiation (+1) site ([\[link\]](#)). This sequence is not identical to the *E. coli* -10 box, but it conserves the A–T rich element. The thermostability of A–T bonds is low and this helps the DNA template to locally unwind in preparation for transcription.

Instead of the simple σ factor that helps bind the prokaryotic RNA polymerase to its promoter, eukaryotes assemble a complex of transcription factors required to recruit RNA polymerase II to a protein coding gene. Transcription factors that bind to the promoter are called *basal transcription factors*. These basal factors are all called TFII (for Transcription Factor/polymerase II) plus an additional letter (A–J). The core complex is TFIID, which includes a TATA-binding protein (TBP). The

other transcription factors systematically fall into place on the DNA template, with each one further stabilizing the pre-initiation complex and contributing to the recruitment of RNA polymerase II.

Visual Connection

A generalized promoter of a gene transcribed by RNA polymerase II is shown. Transcription factors recognize the promoter. RNA polymerase II then binds and forms the transcription initiation complex.



A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacteria to transcribe the gene?

Some eukaryotic promoters also have a conserved **CAAT box** (GGCCAATCT) at approximately -80.

Further upstream of the TATA box, eukaryotic promoters may also contain one or more **GC-rich boxes** (GGCG) or **octamer boxes** (ATTTGCAT). These elements bind cellular factors that increase the efficiency of transcription initiation and are often identified in more “active” genes that are constantly being expressed by the cell.

Basal transcription factors are crucial in the formation of a preinitiation complex on the DNA template that subsequently recruits RNA polymerase II for transcription initiation. The complexity of eukaryotic transcription does not end with the polymerases and promoters. An army of other transcription factors, which bind to upstream enhancers and silencers, also help to regulate the frequency with which pre-mRNA is synthesized from a gene. Enhancers and silencers affect the efficiency of transcription but are not necessary for transcription to proceed.

Promoter Structures for RNA Polymerases I and III

The processes of bringing RNA polymerases I and III to the DNA template involve slightly less complex collections of transcription factors, but the general theme is the same.

The conserved promoter elements for genes transcribed by polymerases I and III differ from

those transcribed by RNA polymerase II. RNA polymerase I transcribes genes that have two GC-rich promoter sequences in the -45 to +20 region. These sequences alone are sufficient for transcription initiation to occur, but promoters with additional sequences in the region from -180 to -105 upstream of the initiation site will further enhance initiation. Genes that are transcribed by RNA polymerase III have upstream promoters or promoters that occur within the genes themselves.

Eukaryotic transcription is a tightly regulated process that requires a variety of proteins to interact with each other and with the DNA strand. Although the process of transcription in eukaryotes involves a greater metabolic investment than in prokaryotes, it ensures that the cell transcribes precisely the pre-mRNAs that it needs for protein synthesis.

Evolution Connection

The Evolution of Promoters

The evolution of genes may be a familiar concept. Mutations can occur in genes during DNA replication, and the result may or may not be beneficial to the cell. By altering an enzyme, structural protein, or some other factor, the process of mutation can transform functions or physical features. However, eukaryotic promoters and other gene regulatory sequences may evolve as well. For

instance, consider a gene that, over many generations, becomes more valuable to the cell. Maybe the gene encodes a structural protein that the cell needs to synthesize in abundance for a certain function. If this is the case, it would be beneficial to the cell for that gene's promoter to recruit transcription factors more efficiently and increase gene expression.

Scientists examining the evolution of promoter sequences have reported varying results. In part, this is because it is difficult to infer exactly where a eukaryotic promoter begins and ends. Some promoters occur within genes; others are located very far upstream, or even downstream, of the genes they are regulating. However, when researchers limited their examination to human core promoter sequences that were defined experimentally as sequences that bind the preinitiation complex, they found that promoters evolve even faster than protein-coding genes.

It is still unclear how promoter evolution might correspond to the evolution of humans or other complex organisms. However, the evolution of a promoter to effectively make more or less of a given gene product is an intriguing alternative to the evolution of the genes themselves.^[footnote]

H Liang et al., "Fast evolution of core promoters in primate genomes," *Molecular Biology and Evolution* 25 (2008): 1239–44.

Eukaryotic Elongation and Termination

Following the formation of the preinitiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed as it does in prokaryotes with the polymerase synthesizing pre-mRNA in the 5' to 3' direction. As discussed previously, RNA polymerase II transcribes the major share of eukaryotic genes, so in this section we will focus on how this polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the DNA template is considerably more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These *DNA–histone complexes*, collectively called nucleosomes, are regularly spaced and include 146 nucleotides of DNA wound around eight histones like thread around a spool.

For polynucleotide synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein complex called **FACT**, which stands for “*facilitates chromatin*

transcription.” This complex pulls histones away from the DNA template as the polymerase moves along it. Once the pre-mRNA is synthesized, the FACT complex replaces the histones to recreate the nucleosomes.

The termination of transcription is different for the different polymerases. Unlike in prokaryotes, elongation by RNA polymerase II in eukaryotes takes place 1,000 to 2,000 nucleotides beyond the end of the gene being transcribed. This pre-mRNA tail is subsequently removed by cleavage during mRNA processing. On the other hand, RNA polymerases I and III require termination signals. Genes transcribed by RNA polymerase I contain a specific 18-nucleotide sequence that is recognized by a termination protein. The process of termination in RNA polymerase III involves an mRNA hairpin similar to rho-independent termination of transcription in prokaryotes.

Section Summary

Transcription in eukaryotes involves one of three types of polymerases, depending on the gene being transcribed. RNA polymerase II transcribes all of the protein-coding genes, whereas RNA polymerase I transcribes the tandemly duplicated rRNA genes, and RNA polymerase III transcribes various small RNAs, like the 5S rRNA, tRNA, and small nuclear

RNA genes. The initiation of transcription in eukaryotes involves the binding of several transcription factors to complex promoter sequences that are usually located upstream of the gene being copied. The mRNA is synthesized in the 5' to 3' direction, and the FACT complex moves and reassembles nucleosomes as the polymerase passes by. Whereas RNA polymerases I and III terminate transcription by protein- or RNA hairpin-dependent methods, RNA polymerase II transcribes for 1,000 or more nucleotides beyond the gene template and cleaves the excess during pre-mRNA processing.

Visual Connection Questions

[\[link\]](#) A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacteria to transcribe the gene?

[\[link\]](#) No. Prokaryotes use different promoters than eukaryotes.

Review Questions

Which feature of promoters can be found in both prokaryotes and eukaryotes?

1. GC box
 2. TATA box
 3. octamer box
 4. -10 and -35 sequences
-

B

What transcripts will be most affected by low levels of α -amanitin?

1. 18S and 28S rRNAs
 2. pre-mRNAs
 3. 5S rRNAs and tRNAs
 4. other small nuclear RNAs
-

B

How do enhancers and promoters differ?

1. Enhancers bind transcription factors to silence gene expression, while promoters activate transcription.
2. Enhancers increase the efficiency of gene expression, but are not essential for transcription. Promoter recognition is

- essential to transcription initiation.
3. Promoters bind transcription factors to increase the efficiency of transcription. Enhancers bind RNA polymerases to initiate transcription.
 4. There is no difference. Both are transcription factor-binding sequences in DNA.
-

B

Critical Thinking Questions

A scientist observes that a cell has an RNA polymerase deficiency that prevents it from making proteins. Describe three additional observations that would together support the conclusion that a defect in RNA polymerase I activity, and not problems with the other polymerases, causes the defect.

To determine that a RNA polymerase I mutation or deficiency is causing the defect in protein production, the scientist would need to make observations that provide evidence that RNA polymerases II and III are working in the cell.

The observations eliminating RNA polymerase II as the defect could include:

- Transcription of mRNAs in the nucleus;
- Presence of processed mRNAs in the cytoplasm

The observations eliminating RNA polymerase III could include:

- Isolation of small nuclear RNAs from the cell;
- Isolation of microRNAs from the cell;
- Transcription of 5S rRNA in the nucleus;
- Presence of tRNAs in the cytoplasm

The observations implicating RNA polymerase I could include:

- A lack of functional ribosomes in the cytoplasm (RNA polymerase I or III);
- A lack of RNA polymerase I protein;
- RNA polymerase I protein is non-functional

Glossary

CAAT box

(GGCCAATCT) essential eukaryotic promoter sequence involved in binding transcription factors

FACT

complex that “facilitates chromatin transcription” by disassembling nucleosomes ahead of a transcribing RNA polymerase II and reassembling them after the polymerase passes by

GC-rich box

(GGCG) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

Octamer box

(ATTTGCAT) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

preinitiation complex

cluster of transcription factors and other proteins that recruit RNA polymerase II for transcription of a DNA template

small nuclear RNA

molecules synthesized by RNA polymerase III that have a variety of functions, including splicing pre-mRNAs and regulating transcription factors

RNA Processing in Eukaryotes

By the end of this section, you will be able to do the following:

- Describe the different steps in RNA processing
- Understand the significance of exons, introns, and splicing for mRNAs
- Explain how tRNAs and rRNAs are processed

After transcription, eukaryotic pre-mRNAs must undergo several processing steps before they can be translated. Eukaryotic (and prokaryotic) tRNAs and rRNAs also undergo processing before they can function as components in the protein-synthesis machinery.

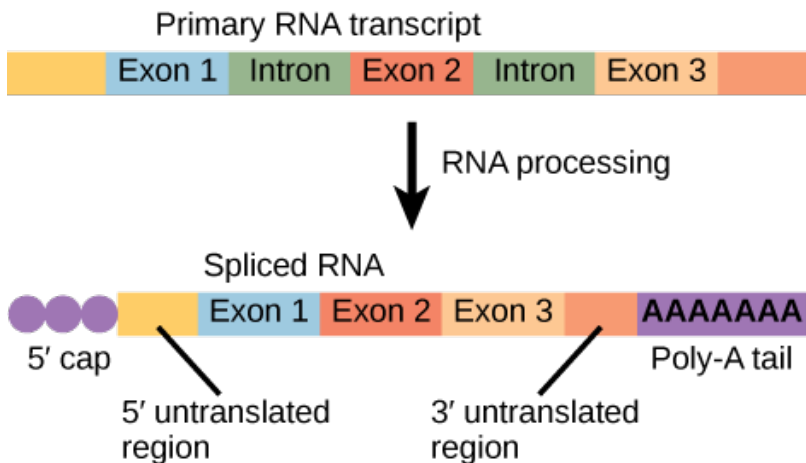
Eukaryotic mRNA contains introns that must be spliced out. A 5' cap and 3' poly-A tail are also added.

mRNA Processing

The eukaryotic pre-mRNA undergoes extensive processing before it is ready to be translated. Eukaryotic protein-coding sequences are not continuous, as they are in prokaryotes. The coding sequences (exons) are interrupted by noncoding introns, which must be removed to make a translatable mRNA. The additional steps involved in eukaryotic mRNA maturation also create a molecule with a much longer half-life than a prokaryotic

mRNA. Eukaryotic mRNAs last for several hours, whereas the typical *E. coli* mRNA lasts no more than five seconds.

Pre-mRNAs are first coated in RNA-stabilizing proteins; these protect the pre-mRNA from degradation while it is processed and exported out of the nucleus. The three most important steps of pre-mRNA processing are the addition of stabilizing and signaling factors at the 5' and 3' ends of the molecule, and the removal of the introns ([\[link\]](#)). In rare cases, the mRNA transcript can be “edited” after it is transcribed.



Evolution Connection

RNA Editing in Trypanosomes

The trypanosomes are a group of protozoa that include the pathogen *Trypanosoma brucei*, which causes nagana in cattle and sleeping sickness in

humans throughout great areas of Africa ([\[link\]](#)).

The trypanosome is carried by biting flies in the genus *Glossina* (commonly called tsetse flies).

Trypanosomes, and virtually all other eukaryotes, have organelles called mitochondria that supply the cell with chemical energy. Mitochondria are organelles that express their own DNA and are believed to be the remnants of a symbiotic relationship between a eukaryote and an engulfed prokaryote. The mitochondrial DNA of trypanosomes exhibit an interesting exception to the central dogma: their pre-mRNAs do not have the correct information to specify a functional protein. Usually, this is because the mRNA is missing several U nucleotides. The cell performs an additional RNA processing step called **RNA editing** to remedy this.

Trypanosoma brucei is the causative agent of sleeping sickness in humans. The mRNAs of this pathogen must be modified by the addition of nucleotides before protein synthesis can occur.

(credit: modification of work by Torsten Ochsenreiter)



Other genes in the mitochondrial genome encode 40- to 80-nucleotide guide RNAs. One or more of these molecules interacts by complementary base pairing with some of the nucleotides in the pre-mRNA transcript. However, the *guide RNA* has more A nucleotides than the pre-mRNA has U nucleotides with which to bind. In these regions, the guide RNA loops out. The 3' ends of guide RNAs have a long poly-U tail, and these U bases are inserted in regions of the pre-mRNA transcript at which the guide RNAs are looped. This process is entirely mediated by RNA molecules. That is, guide RNAs—rather than proteins—serve as the catalysts in RNA editing.

RNA editing is not just a phenomenon of trypanosomes. In the mitochondria of some plants,

almost all pre-mRNAs are edited. RNA editing has also been identified in mammals such as rats, rabbits, and even humans. What could be the evolutionary reason for this additional step in pre-mRNA processing? One possibility is that the mitochondria, being remnants of ancient prokaryotes, have an equally ancient RNA-based method for regulating gene expression. In support of this hypothesis, edits made to pre-mRNAs differ depending on cellular conditions. Although speculative, the process of RNA editing may be a holdover from a primordial time when RNA molecules, instead of proteins, were responsible for catalyzing reactions.

5' Capping

While the pre-mRNA is still being synthesized, a **7-methylguanosine cap** is added to the 5' end of the growing transcript by a phosphate linkage. This functional group protects the nascent mRNA from degradation. In addition, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

3' Poly-A Tail

Once elongation is complete, the pre-mRNA is cleaved by an endonuclease between an AAUAAA

consensus sequence and a GU-rich sequence, leaving the AAUAAA sequence on the pre-mRNA. An enzyme called poly-A polymerase then adds a string of approximately 200 A residues, called the **poly-A tail**. This modification further protects the pre-mRNA from degradation and is also the binding site for a protein necessary for exporting the processed mRNA to the cytoplasm.

Pre-mRNA Splicing

Eukaryotic genes are composed of **exons**, which correspond to protein-coding sequences (*ex-on* signifies that they are *expressed*), and *intervening* sequences called **introns** (*int-ron* denotes their *intervening* role), which may be involved in gene regulation but are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins.

The discovery of introns came as a surprise to researchers in the 1970s who expected that pre-mRNAs would specify protein sequences without further processing, as they had observed in prokaryotes. The genes of higher eukaryotes very often contain one or more introns. These regions may correspond to regulatory sequences; however, the biological significance of having many introns or having very long introns in a gene is unclear. It is possible that introns slow down gene expression because it takes longer to transcribe pre-mRNAs

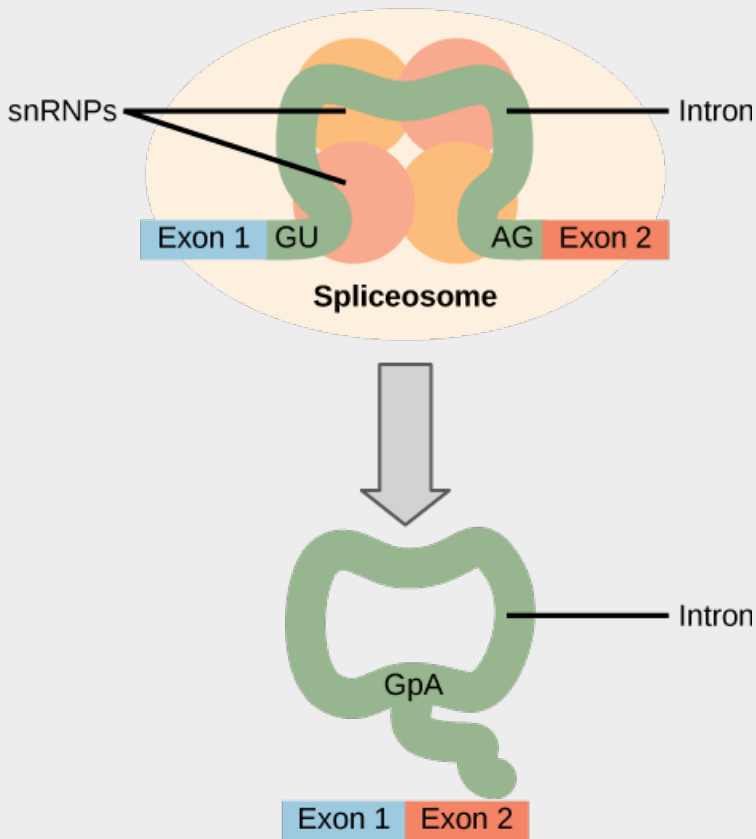
with lots of introns. Alternatively, introns may be nonfunctional sequence remnants left over from the fusion of ancient genes throughout the course of evolution. This is supported by the fact that separate exons often encode separate protein subunits or domains. For the most part, the sequences of introns can be mutated without ultimately affecting the protein product.

All of a pre-mRNA's introns must be completely and precisely removed before protein synthesis. If the process errs by even a single nucleotide, the reading frame of the rejoined exons would shift, and the resulting protein would be dysfunctional. The process of removing introns and reconnecting exons is called **splicing** ([\[link\]](#)). Introns are removed and degraded while the pre-mRNA is still in the nucleus. Splicing occurs by a sequence-specific mechanism that ensures introns will be removed and exons rejoined with the accuracy and precision of a single nucleotide. Although the intron itself is noncoding, the beginning and end of each intron is marked with specific nucleotides: GU at the 5' end and AG at the 3' end of the intron. The splicing of pre-mRNAs is conducted by complexes of proteins and RNA molecules called spliceosomes.

Visual Connection

Pre-mRNA splicing involves the precise removal of

introns from the primary RNA transcript. The splicing process is catalyzed by protein complexes called spliceosomes that are composed of proteins and RNA molecules called small nuclear RNAs (snRNAs). Spliceosomes recognize sequences at the 5' and 3' end of the intron.



Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

Note that more than 70 individual introns can be present, and each has to undergo the process of splicing—in addition to 5' capping and the addition of a poly-A tail—just to generate a single, translatable mRNA molecule.

Link to Learning

See how introns are removed during RNA splicing [at this website](#).

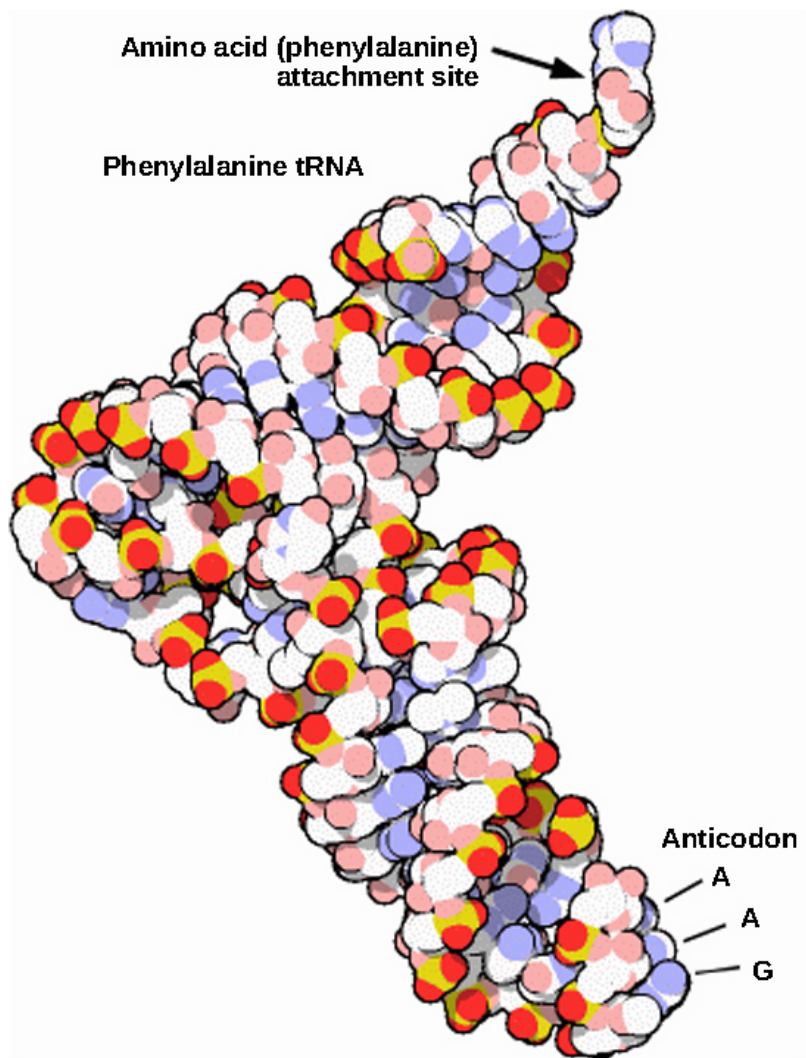
This is a space-filling model of a tRNA molecule that adds the amino acid phenylalanine to a growing polypeptide chain. The anticodon AAG binds the Codon UUC on the mRNA. The amino acid phenylalanine is attached to the other end of the tRNA.

Processing of tRNAs and rRNAs

The tRNAs and rRNAs are structural molecules that have roles in protein synthesis; however, these RNAs are not themselves translated. Pre-rRNAs are transcribed, processed, and assembled into ribosomes in the nucleolus. Pre-tRNAs are transcribed and processed in the nucleus and then released into the cytoplasm where they are linked to free amino acids for protein synthesis.

Most of the tRNAs and rRNAs in eukaryotes and prokaryotes are first transcribed as a long precursor molecule that spans multiple rRNAs or tRNAs. Enzymes then cleave the precursors into subunits corresponding to each structural RNA. Some of the bases of pre-rRNAs are *methyalted*; that is, a $-CH_3$ methyl functional group is added for stability. Pre-tRNA molecules also undergo methylation. As with pre-mRNAs, subunit excision occurs in eukaryotic pre-RNAs destined to become tRNAs or rRNAs.

Mature rRNAs make up approximately 50 percent of each ribosome. Some of a ribosome's RNA molecules are purely structural, whereas others have catalytic or binding activities. Mature tRNAs take on a three-dimensional structure through local regions of base pairing stabilized by intramolecular hydrogen bonding. The tRNA folds to position the amino acid binding site at one end and the **anticodon** at the other end ([\[link\]](#)). The anticodon is a three-nucleotide sequence in a tRNA that interacts with an mRNA codon through complementary base pairing.



Section Summary

Eukaryotic pre-mRNAs are modified with a 5' methylguanosine cap and a poly-A tail. These structures protect the mature mRNA from

degradation and help export it from the nucleus. Pre-mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. Only finished mRNAs that have undergone 5' capping, 3' polyadenylation, and intron splicing are exported from the nucleus to the cytoplasm. Pre-rRNAs and pre-tRNAs may be processed by intramolecular cleavage, splicing, methylation, and chemical conversion of nucleotides. Rarely, RNA editing is also performed to insert missing bases after an mRNA has been synthesized.

Visual Connection Questions

[\[link\]](#) Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

[\[link\]](#) Mutations in the spliceosome recognition sequence at each end of the intron, or in the proteins and RNAs that make up the spliceosome, may impair splicing. Mutations may also add new spliceosome recognition sites. Splicing errors could lead to introns being

retained in spliced RNA, exons being excised, or changes in the location of the splice site.

Review Questions

Which pre-mRNA processing step is important for initiating translation?

1. poly-A tail
2. RNA editing
3. splicing
4. 7-methylguanosine cap

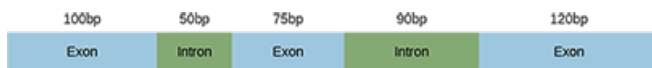
D

What processing step enhances the stability of pre-tRNAs and pre-rRNAs?

1. methylation
2. nucleotide modification
3. cleavage
4. splicing

A

A scientist identifies a pre-mRNA with the following structure.



What is the predicted size of the corresponding mature mRNA in base pairs (bp), excluding the 5' cap and 3' poly-A tail?

1. 220bp
2. 295bp
3. 140bp
4. 435bp

B

Critical Thinking Questions

Chronic lymphocytic leukemia patients often harbor nonsense mutations in their spliceosome machinery. Describe how this mutation of the spliceosome would change the final location and sequence of a pre-mRNA.

Nonsense spliceosome mutations would

eliminate the splicing step of mRNA processing, so the mature mRNAs would retain their introns and be perfectly complementary to the entire DNA template sequence. However, the mRNAs would still undergo addition of the 5' cap and poly-A tail, and therefore each has the potential to be exported to the cytoplasm for translation.

Glossary

7-methylguanosine cap

modification added to the 5' end of pre-mRNAs to protect mRNA from degradation and assist translation

anticodon

three-nucleotide sequence in a tRNA molecule that corresponds to an mRNA codon

exon

sequence present in protein-coding mRNA after completion of pre-mRNA splicing

intron

non-protein-coding intervening sequences that are spliced from mRNA during processing

poly-A tail

modification added to the 3' end of pre-mRNAs to protect mRNA from degradation and assist mRNA export from the nucleus

RNA editing

direct alteration of one or more nucleotides in an mRNA that has already been synthesized

splicing

process of removing introns and reconnecting exons in a pre-mRNA

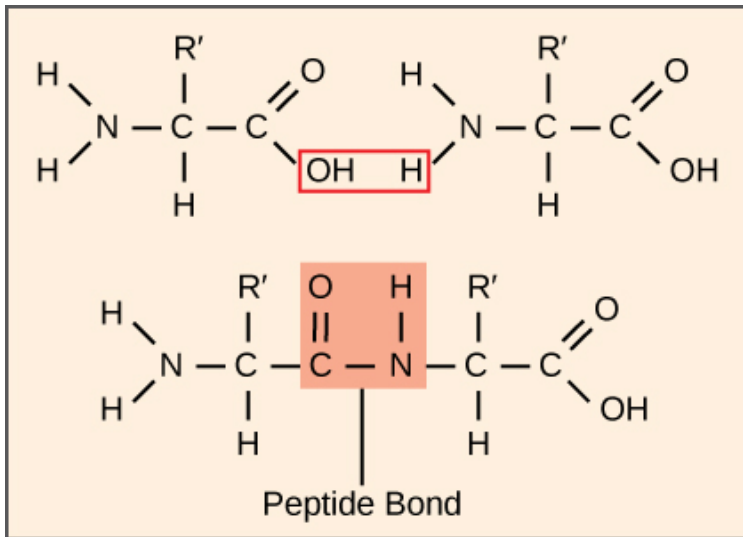
Ribosomes and Protein Synthesis

By the end of this section, you will be able to do the following:

- Describe the different steps in protein synthesis
- Discuss the role of ribosomes in protein synthesis

The synthesis of proteins consumes more of a cell's energy than any other metabolic process. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform virtually every function of a cell. The process of translation, or protein synthesis, involves the decoding of an mRNA message into a polypeptide product. Amino acids are covalently strung together by interlinking peptide bonds in lengths ranging from approximately 50 to more than 1000 amino acid residues. Each individual amino acid has an amino group (NH₂) and a carboxyl (COOH) group. Polypeptides are formed when the amino group of one amino acid forms an amide (i.e., peptide) bond with the carboxyl group of another amino acid ([\[link\]](#)). This reaction is catalyzed by ribosomes and generates one water molecule. A peptide bond links the carboxyl end of one amino acid with the amino end of another, producing one water molecule during the process. For simplicity in this image, only the functional groups involved in the peptide bond are shown. The R and R'

designations refer to the rest of each amino acid structure.



The Protein Synthesis Machinery

In addition to the mRNA template, many molecules and macromolecules contribute to the process of translation. The composition of each component may vary across species; for example, ribosomes may consist of different numbers of rRNAs and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors. (Note: A ribosome can be thought of as an enzyme whose amino acid binding sites are specified by mRNA.)

Link to Learning

Click through the steps of this [PBS interactive](#) to see protein synthesis in action.

Ribosomes

Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. In *E. coli*, there are between 10,000 and 70,000 ribosomes present in each cell at any given time. A **ribosome** is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs.

Ribosomes exist in the cytoplasm of prokaryotes and in the cytoplasm and rough endoplasmic reticulum of eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to prokaryotic ribosomes (and have similar drug sensitivities) than the ribosomes just outside their outer membranes in the cytoplasm. Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the initiation of translation. *In E. coli, the small subunit is described as 30S, and the large subunit is 50S, for a total of 70S (recall that Svedberg units are not additive).*

Mammalian ribosomes have a small 40S subunit and a large 60S subunit, for a total of 80S. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds tRNAs. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5' to 3' and synthesizing the polypeptide from the N terminus to the C terminus. The complete mRNA/poly-ribosome structure is called a **polysome**.

tRNAs

The tRNAs are structural RNA molecules that were transcribed from genes by RNA polymerase III. Depending on the species, 40 to 60 types of tRNAs exist in the cytoplasm. Transfer RNAs serve as adaptor molecules. Each tRNA carries a specific amino acid and recognizes one or more of the mRNA codons that define the order of amino acids in a protein. Aminoacyl-tRNAs bind to the ribosome and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually “translate” the language of RNA into the language of proteins.

Of the 64 possible mRNA codons—or triplet combinations of A, U, G, and C—three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of

these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA anticodon can base pair with one or more of the mRNA codons for its amino acid. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind a leucine tRNA expressing the complementary sequence, GAU. The ability of some tRNAs to match more than one codon is what gives the genetic code its blocky structure.

As the adaptor molecules of translation, it is surprising that tRNAs can fit so much specificity into such a small package. Consider that tRNAs need to interact with three factors: 1) they must be recognized by the correct aminoacyl synthetase (see below); 2) they must be recognized by ribosomes; and 3) they must bind to the correct sequence in mRNA.

Aminoacyl tRNA Synthetases

The process of pre-tRNA synthesis by RNA polymerase III only creates the RNA portion of the adaptor molecule. The corresponding amino acid must be added later, once the tRNA is processed and exported to the cytoplasm. Through the process of tRNA “charging,” each tRNA molecule is linked to its correct amino acid by one of a group of enzymes called **aminoacyl tRNA synthetases**. At least one type of aminoacyl tRNA synthetase exists for each of

the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes first bind and hydrolyze ATP to catalyze a high-energy bond between an amino acid and adenosine monophosphate (AMP); a pyrophosphate molecule is expelled in this reaction. The activated amino acid is then transferred to the tRNA, and AMP is released. The term "charging" is appropriate, since the high-energy bond that attaches an amino acid to its tRNA is later used to drive the formation of the peptide bond. Each tRNA is named for its amino acid.

The Mechanism of Protein Synthesis

As with mRNA synthesis, protein synthesis can be divided into three phases: *initiation*, *elongation*, and *termination*. The process of translation is similar in prokaryotes and eukaryotes. Here we'll explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

Initiation of Translation

Protein synthesis begins with the formation of an **initiation complex**. In *E. coli*, this complex involves the small 30S ribosome, the mRNA template, three initiation factors (IFs; IF-1, IF-2, and IF-3), and a

special **initiator tRNA**, called tRNA_{Metf}.

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the Shine-Dalgarno sequence (AGGAGG), interacts with the rRNA molecules that compose the ribosome. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. Guanosine triphosphate (GTP), which is a purine nucleotide triphosphate, acts as an energy source during translation—both at the start of elongation and during the ribosome's translocation. Binding of the mRNA to the 30S ribosome also requires IF-III.

The initiator tRNA then interacts with the **start codon** AUG (or rarely, GUG). This tRNA carries the amino acid methionine, which is formylated after its attachment to the tRNA. The formylation creates a "faux" peptide bond between the formyl carboxyl group and the amino group of the methionine. Binding of the fMet-tRNA_{Metf} is mediated by the initiation factor IF-2. The fMet begins every polypeptide chain synthesized by *E. coli*, but it is usually removed after translation is complete. When an in-frame AUG is encountered during translation elongation, a non-formylated methionine is inserted by a regular Met-tRNA_{Met}. After the formation of the initiation complex, the 30S ribosomal subunit is joined by the 50S subunit to form the translation complex. In eukaryotes, a similar initiation complex forms, comprising mRNA, the 40S small ribosomal

subunit, eukaryotic IFs, and nucleoside triphosphates (GTP and ATP). The methionine on the charged initiator tRNA, called Met-tRNA_i, is not formylated. However, Met-tRNA_i is distinct from other Met-tRNAs in that it can bind IFs.

Instead of depositing at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5' end of the mRNA. A cap-binding protein (CBP) and several other IFs assist the movement of the ribosome to the 5' cap. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. According to **Kozak's rules**, the nucleotides around the AUG indicate whether it is the correct start codon. Kozak's rules state that the following consensus sequence must appear around the AUG of vertebrate genes: 5'-gccRccAUGG-3'. The R (for purine) indicates a site that can be either A or G, but cannot be C or U. Essentially, the closer the sequence is to this consensus, the higher the efficiency of translation.

Once the appropriate AUG is identified, the other proteins and CBP dissociate, and the 60S subunit binds to the complex of Met-tRNA_i, mRNA, and the 40S subunit. This step completes the initiation of translation in eukaryotes.

Translation, Elongation, and Termination

In prokaryotes and eukaryotes, the basics of elongation are the same, so we will review elongation from the perspective of *E. coli*. When the translation complex is formed, the tRNA binding region of the ribosome consists of three compartments. The A (aminoacyl) site binds incoming charged aminoacyl tRNAs. The P (peptidyl) site binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The E (exit) site releases dissociated tRNAs so that they can be recharged with free amino acids. The initiating methionyl-tRNA, however, occupies the P site at the beginning of the elongation phase of translation in both prokaryotes and eukaryotes.

During translation elongation, the mRNA template provides tRNA binding specificity. As the ribosome moves along the mRNA, each mRNA codon comes into register, and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically and randomly (?).

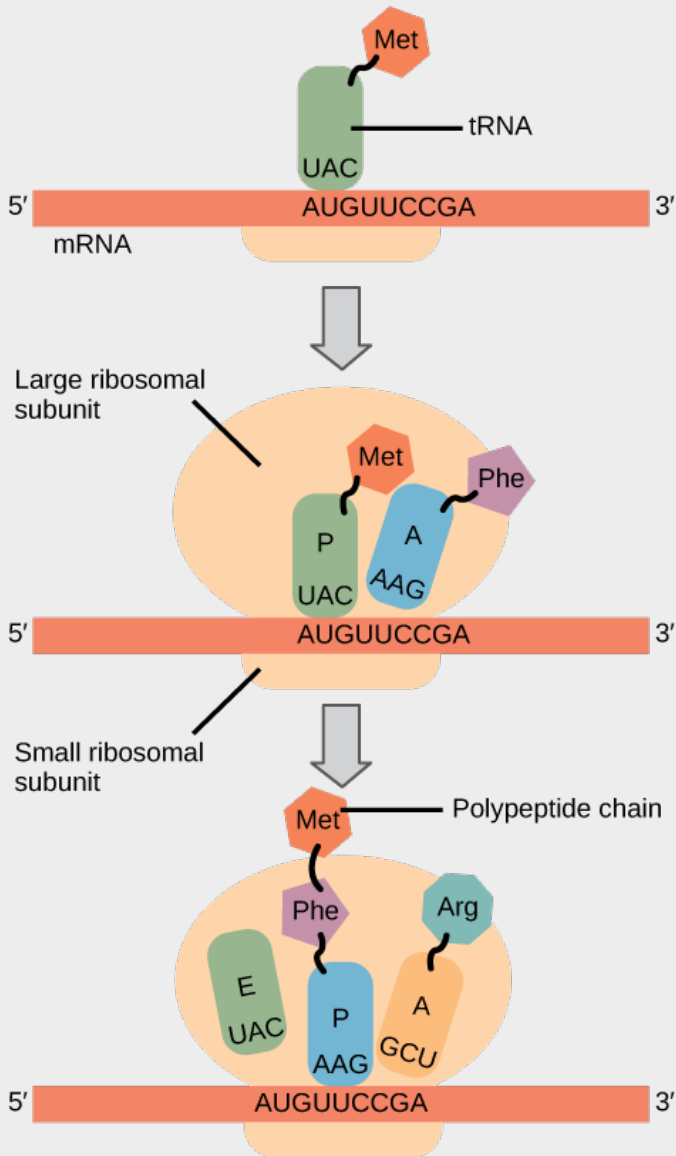
Elongation proceeds with charged tRNAs sequentially entering and leaving the ribosome as each new amino acid is added to the polypeptide

chain. Movement of a tRNA from A to P to E site is induced by conformational changes that advance the ribosome by three bases in the 3' direction. The energy for each step along the ribosome is donated by elongation factors that hydrolyze GTP. GTP energy is required both for the binding of a new aminoacyl-tRNA to the A site and for its translocation to the P site after formation of the peptide bond. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, an RNA-based enzyme that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation is derived from the high-energy bond linking each amino acid to its tRNA. After peptide bond formation, the A-site tRNA that now holds the growing peptide chain moves to the P site, and the P-site tRNA that is now empty moves to the E site and is expelled from the ribosome ([\[link\]](#)). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino-acid protein can be translated in just 10 seconds.

Visual Connection

Translation begins when an initiator tRNA anticodon recognizes a start codon on mRNA

bound to a small ribosomal subunit. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, successive tRNAs move through the ribosome and the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.



Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein

synthesis?

Tetracycline would directly affect:

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

Chloramphenicol would directly affect:

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

Termination of translation occurs when a nonsense codon (UAA, UAG, or UGA) is encountered. Upon aligning with the A site, these nonsense codons are recognized by *protein release factors* that resemble tRNAs. The releasing factors in both prokaryotes and eukaryotes instruct peptidyl transferase to add a water molecule to the carboxyl end of the P-site amino acid. This reaction forces the P-site amino acid to detach from its tRNA, and the newly made protein is released. The small and large ribosomal subunits dissociate from the mRNA and from each other; they are recruited almost immediately into another translation initiation complex. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

Protein Folding, Modification, and Targeting

During and after translation, individual amino acids may be chemically modified, signal sequences appended, and the new protein “folded” into a distinct three-dimensional structure as a result of intramolecular interactions. A **signal sequence** is a short sequence at the amino end of a protein that directs it to a specific cellular compartment. These sequences can be thought of as the protein’s “train ticket” to its ultimate destination, and are recognized by signal-recognition proteins that act as conductors. For instance, a specific signal sequence terminus will direct a protein to the mitochondria or chloroplasts (in plants). Once the protein reaches its cellular destination, the signal sequence is usually clipped off.

Many proteins fold spontaneously, but some proteins require helper molecules, called *chaperones*, to prevent them from aggregating during the complicated process of folding. Even if a protein is properly specified by its corresponding mRNA, it could take on a completely dysfunctional shape if abnormal temperature or pH conditions prevent it from folding correctly.

Section Summary

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit binds to the mRNA template either at the Shine-Dalgarno sequence (prokaryotes) or the 5' cap (eukaryotes). Translation begins at the initiating AUG on the mRNA, specifying methionine. The formation of peptide bonds occurs between sequential amino acids matched to the mRNA template by their tRNAs according to the genetic code. Charged tRNAs enter the ribosomal A site, and their amino acid bonds with the amino acid at the P site. The entire mRNA is translated in three-nucleotide “steps” of the ribosome. When a nonsense codon is encountered, a release factor binds and dissociates the components and frees the new protein. Folding of the protein occurs during and after translation.

Visual Connection Questions

[\[link\]](#) Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

Tetracycline would directly affect:

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

Chloramphenicol would directly affect

1. tRNA binding to the ribosome
2. ribosome assembly
3. growth of the protein chain

[\[link\]](#) Tetracycline: a; Chloramphenicol: c.

Review Questions

The RNA components of ribosomes are synthesized in the _____.

1. cytoplasm
2. nucleus
3. nucleolus
4. endoplasmic reticulum

In any given species, there are at least how many types of aminoacyl tRNA synthetases?

1. 20
2. 40
3. 100
4. 200

A

A scientist introduces a mutation that makes the 60S ribosomal subunit nonfunctional in a human cell line. What would be the predicted effect on translation?

1. Translation stalls after the initiation AUG codon is identified.
2. The ribosome cannot catalyze the formation of peptide bonds between the tRNAs in the A and P sites.
3. The ribosome cannot interact with mRNAs.
4. tRNAs cannot exit the E site of the ribosome.

A

Critical Thinking Questions

Transcribe and translate the following DNA sequence (nontemplate strand): 5'-ATGGCCGGTTATTAAGCA-3'

The mRNA would be: 5'-AUGGCCGGUUAUUAAGCA-3'. The protein would be: MAGY. Even though there are six codons, the fifth codon corresponds to a stop, so the sixth codon would not be translated.

Explain how single nucleotide changes can have vastly different effects on protein function.

Nucleotide changes in the third position of codons may not change the amino acid and would have no effect on the protein. Other nucleotide changes that change important amino acids or create or delete start or stop codons would have severe effects on the amino acid sequence of the protein.

A normal mRNA that reads 5' – UGCCAUGGUAAUAACACAUGAGGCCUGAAC– 3' has an insertion mutation that changes the sequence to 5' - UGCCAUGGUUAAUAACACAUGAGGCCUGAAC– 3'. Translate the original mRNA and the

mutated mRNA, and explain how insertion mutations can have dramatic effects on proteins. (Hint: Be sure to find the initiation site.)

Original mRNA: 5' –UGCC AUG GUA AUA ACA CAU GAG GCC UGA AC– 3';

Translation: Met – Val – Ile – Thr – His – Glu – Ala;

Mutated mRNA: 5' –UGCC AUG GUU AAU AAC ACA UGA GGCCUGAAC– 3';

Translation: Met – Val – Asn – Asn – Thr;

Insertion mutations can have dramatic effects on proteins because they shift the reading frame for the codons. This changes the amino acids encoded by the mRNA, and can introduce premature start or stop sites.

Glossary

aminoacyl tRNA synthetase

enzyme that “charges” tRNA molecules by catalyzing a bond between the tRNA and a corresponding amino acid

initiator tRNA

in prokaryotes, called tRNA^{fMet}; in eukaryotes, called tRNAⁱ; a tRNA that interacts with a start codon, binds directly to the ribosome P site, and links to a special methionine to begin a polypeptide chain

Kozak's rules

determines the correct initiation AUG in a eukaryotic mRNA; the following consensus sequence must appear around the AUG: 5'-GCC(**purine**)CCAUGG-3'; the bolded bases are most important

peptidyl transferase

RNA-based enzyme that is integrated into the 50S ribosomal subunit and catalyzes the formation of peptide bonds

polysome

mRNA molecule simultaneously being translated by many ribosomes all going in the same direction

Shine-Dalgarno sequence

(AGGAGG); initiates prokaryotic translation by interacting with rRNA molecules comprising the 30S ribosome

signal sequence

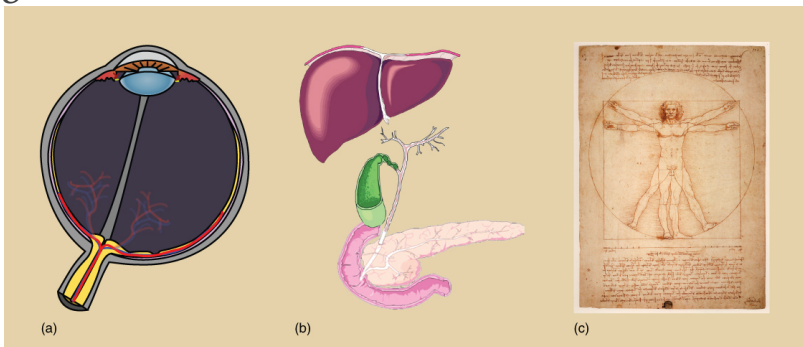
short tail of amino acids that directs a protein to a specific cellular compartment

start codon

AUG (or rarely, GUG) on an mRNA from which translation begins; always specifies methionine

Introduction

class = "introduction" The genetic content of each somatic cell in an organism is the same, but not all genes are expressed in every cell. The control of which genes are expressed dictates whether a cell is, for example, (a) an eye cell or (b) a liver cell. It is the differential gene expression patterns that arise in different cells that give rise to (c) a complete organism.



Each somatic cell in the body generally contains the same DNA. A few exceptions include red blood cells, which contain no DNA in their mature state, and some immune system cells that rearrange their DNA while producing antibodies. In general, however, the genes that determine whether you have green eyes, brown hair, and how fast you metabolize food are the same in the cells in your eyes and your liver, even though these organs function quite differently. If each cell has the same DNA, how is it that cells or organs are different? Why do cells in the eye differ so dramatically from cells in the liver?

Whereas each cell shares the same genome and DNA sequence, each cell does not turn on, or express, the same set of genes. Each cell type needs a different set of proteins to perform its function. Therefore, only a small subset of proteins is expressed in a cell. For the proteins to be expressed, the DNA must be transcribed into RNA and the RNA must be translated into protein. In a given cell type, not all genes encoded in the DNA are transcribed into RNA or translated into protein because specific cells in our body have specific functions. Specialized proteins that make up the eye (iris, lens, and cornea) are only expressed in the eye, whereas the specialized proteins in the heart (pacemaker cells, heart muscle, and valves) are only expressed in the heart. At any given time, only a subset of all of the genes encoded by our DNA are expressed and translated into proteins. The expression of specific genes is a highly regulated process with many levels and stages of control. This complexity ensures the proper expression in the proper cell at the proper time.

Regulation of Gene Expression

By the end of this section, you will be able to do the following:

- Discuss why every cell does not express all of its genes all of the time
- Describe how prokaryotic gene regulation occurs at the transcriptional level
- Discuss how eukaryotic gene regulation occurs at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels

For a cell to function properly, necessary proteins must be synthesized at the proper time and place. All cells control or regulate the synthesis of proteins from information encoded in their DNA. The process of turning on a gene to produce RNA and protein is called **gene expression**. Whether in a simple unicellular organism or a complex multi-cellular organism, each cell controls when and how its genes are expressed. For this to occur, there must be internal chemical mechanisms that control when a gene is expressed to make RNA and protein, how much of the protein is made, and when it is time to stop making that protein because it is no longer needed.

The regulation of gene expression conserves energy and space. It would require a significant amount of energy for an organism to express every gene at all

times, so it is more energy efficient to turn on the genes only when they are required. In addition, only expressing a subset of genes in each cell saves space because DNA must be unwound from its tightly coiled structure to transcribe and translate the DNA. Cells would have to be enormous if every protein were expressed in every cell all the time.

The control of gene expression is extremely complex. Malfunctions in this process are detrimental to the cell and can lead to the development of many diseases, including cancer.

Regulation in prokaryotes and eukaryotes. Prokaryotic transcription and translation occur simultaneously in the cytoplasm, and regulation occurs at the transcriptional level. Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, and during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.

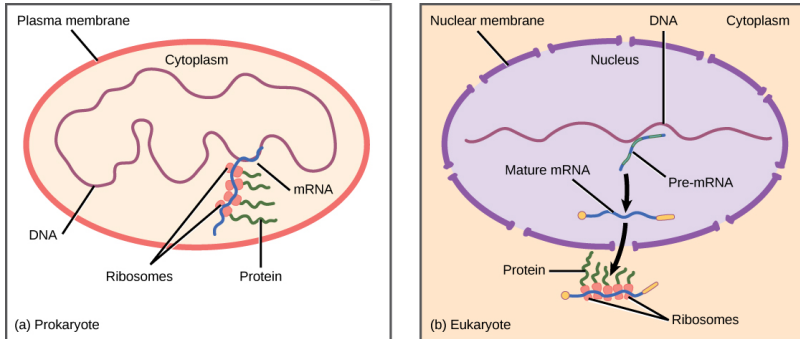
Prokaryotic versus Eukaryotic Gene Expression

To understand how gene expression is regulated, we must first understand how a gene codes for a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different manners.

Prokaryotic organisms are single-celled organisms that lack a cell nucleus, and their DNA therefore floats freely in the cell cytoplasm. To synthesize a protein, the processes of transcription and translation occur almost simultaneously. When the resulting protein is no longer needed, transcription stops. As a result, the primary method to control what type of protein and how much of each protein is expressed in a prokaryotic cell is the regulation of DNA transcription. All of the subsequent steps occur automatically. When more protein is required, more transcription occurs. Therefore, in prokaryotic cells, the control of gene expression is mostly at the transcriptional level.

Eukaryotic cells, in contrast, have intracellular organelles that add to their complexity. In eukaryotic cells, the DNA is contained inside the cell's nucleus and there it is transcribed into RNA. The newly synthesized RNA is then transported out of the nucleus into the cytoplasm, where ribosomes translate the RNA into protein. The processes of transcription and translation are *physically separated* by the nuclear membrane; transcription occurs only within the nucleus, and translation occurs only outside the nucleus in the cytoplasm. The regulation of gene expression can occur at all stages of the process ([\[link\]](#)). Regulation may occur when the DNA is uncoiled and loosened from nucleosomes to bind transcription factors (**epigenetic** level), when the RNA is transcribed (**transcriptional** level),

when the RNA is processed and exported to the cytoplasm after it is transcribed (**post-transcriptional level**), when the RNA is translated into protein (**translational level**), or after the protein has been made (**post-translational level**).



The differences in the regulation of gene expression between prokaryotes and eukaryotes are summarized in [\[link\]](#). The regulation of gene expression is discussed in detail in subsequent modules.

Differences in the Regulation of Gene Expression of Prokaryotic and Eukaryotic Organisms

Prokaryotic organisms
Lack a membrane-bound

Eukaryotic organisms
Contain nucleus

nucleus

DNA is found in the cytoplasm

RNA transcription and protein formation occur almost simultaneously

DNA is confined to the nuclear compartment

RNA transcription occurs prior to protein formation, and it takes place in the nucleus. Translation of RNA to protein occurs in the cytoplasm.

Gene expression is regulated primarily at the transcriptional level

Gene expression is regulated at many levels (epigenetic, transcriptional, nuclear shuttling, post-transcriptional, translational, and post-translational)

Evolution Connection

Evolution of Gene Regulation

Prokaryotic cells can only regulate gene expression by controlling the amount of transcription. As eukaryotic cells evolved, the complexity of the control of gene expression increased. For example, with the evolution of eukaryotic cells came compartmentalization of important cellular components and cellular processes. A nuclear region that contains the DNA was formed.

Transcription and translation were physically separated into two different cellular compartments. It therefore became possible to control gene expression by regulating transcription in the nucleus, and also by controlling the RNA levels and protein translation present outside the nucleus. Most gene regulation is done to conserve cell resources. However, other regulatory processes may be defensive. Cellular processes such as developed to protect the cell from viral or parasitic infections. If the cell could quickly shut off gene expression for a short period of time, it would be able to survive an infection when other organisms could not. Therefore, the organism evolved a new process that helped it survive, and it was able to pass this new development to offspring.

Section Summary

While all somatic cells within an organism contain the same DNA, not all cells within that organism express the same proteins. Prokaryotic organisms express most of their genes most of the time.

However, some genes are expressed only when they are needed. Eukaryotic organisms, on the other hand, express only a subset of their genes in any given cell. To express a protein, the DNA is first

transcribed into RNA, which is then translated into proteins, which are then targeted to specific cellular locations. In prokaryotic cells, transcription and translation occur almost simultaneously. In eukaryotic cells, transcription occurs in the nucleus and is separate from the translation that occurs in the cytoplasm. Gene expression in prokaryotes is mostly regulated at the transcriptional level (some epigenetic and post-translational regulation is also present), whereas in eukaryotic cells, gene expression is regulated at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels.

Review Questions

Control of gene expression in eukaryotic cells occurs at which level(s)?

1. only the transcriptional level
2. epigenetic and transcriptional levels
3. epigenetic, transcriptional, and translational levels
4. epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels

Post-translational control refers to:

1. regulation of gene expression after transcription
2. regulation of gene expression after translation
3. control of epigenetic activation
4. period between transcription and translation

B

How does the regulation of gene expression support continued evolution of more complex organisms?

1. Cells can become specialized within a multicellular organism.
2. Organisms can conserve energy and resources.
3. Cells grow larger to accommodate protein production.
4. Both A and B.

D

Critical Thinking Questions

Name two differences between prokaryotic and eukaryotic cells and how these differences benefit multicellular organisms.

Eukaryotic cells have a nucleus, whereas prokaryotic cells do not. In eukaryotic cells, DNA is confined within the nuclear region. Because of this, transcription and translation are physically separated. This creates a more complex mechanism for the control of gene expression that benefits multicellular organisms because it compartmentalizes gene regulation.

Gene expression occurs at many stages in eukaryotic cells, whereas in prokaryotic cells, control of gene expression only occurs at the transcriptional level. This allows for greater control of gene expression in eukaryotes and more complex systems to be developed. Because of this, different cell types can arise in an individual organism.

Describe how controlling gene expression will alter the overall protein levels in the cell.

The cell controls which proteins are expressed and to what level each protein is expressed in the cell. Prokaryotic cells alter the transcription rate to turn genes on or off. This method will increase or decrease protein levels in response to what is needed by the cell. Eukaryotic cells change the accessibility (epigenetic), transcription, or translation of a gene. This will alter the amount of RNA and the lifespan of the RNA to alter the amount of protein that exists. Eukaryotic cells also control protein translation to increase or decrease the overall levels. Eukaryotic organisms are much more complex and can manipulate protein levels by changing many stages in the process.

Glossary

epigenetic

heritable changes that do not involve changes in the DNA sequence

gene expression

processes that control the turning on or turning off of a gene

post-transcriptional

control of gene expression after the RNA molecule has been created but before it is translated into protein

post-translational

control of gene expression after a protein has
been created

Prokaryotic Gene Regulation

By the end of this section, you will be able to do the following:

- Describe the steps involved in prokaryotic gene regulation
- Explain the roles of activators, inducers, and repressors in gene regulation

The DNA of prokaryotes is organized into a circular chromosome, supercoiled within the nucleoid region of the cell cytoplasm. Proteins that are needed for a specific function, or that are involved in the same biochemical pathway, are encoded together in blocks called **operons**. For example, all of the genes needed to use lactose as an energy source are coded next to each other in the lactose (or *lac*) operon, and transcribed into a single mRNA.

In prokaryotic cells, there are three types of regulatory molecules that can affect the expression of operons: repressors, activators, and inducers. Repressors and activators are proteins produced in the cell. Both repressors and activators regulate gene expression by binding to specific DNA sites *adjacent* to the genes they control. *In general, activators bind to the promoter site, while repressors bind to operator regions.* **Repressors** prevent transcription of a gene in response to an external stimulus, whereas **activators** increase the transcription of a gene in response to an external

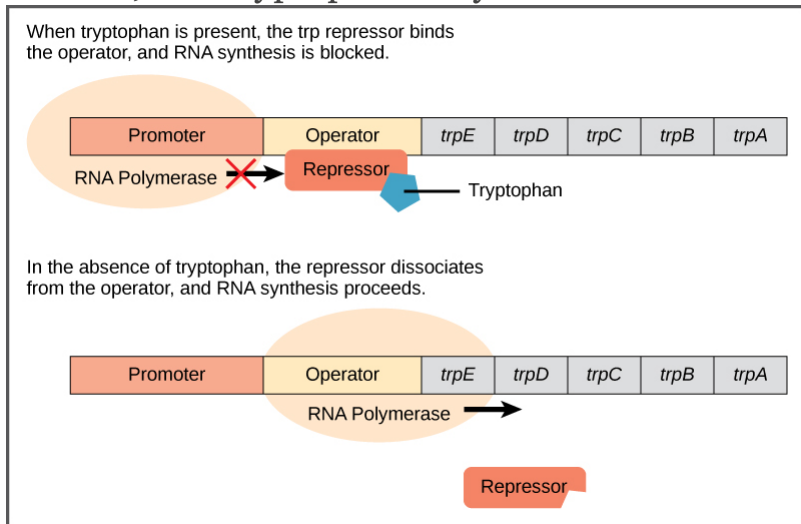
stimulus. Inducers are small molecules that may be produced by the cell or that are in the cell's environment. Inducers either activate or repress transcription depending on the needs of the cell and the availability of substrate.

The tryptophan operon. The five genes that are needed to synthesize tryptophan in *E. coli* are located next to each other in the *trp* operon. When tryptophan is plentiful, two tryptophan molecules bind the repressor protein at the operator sequence. This physically blocks the RNA polymerase from transcribing the tryptophan genes. When tryptophan is absent, the repressor protein does not bind to the operator and the genes are transcribed.

The *trp* Operon: A Repressible Operon

Bacteria such as *Escherichia coli* need amino acids to survive, and are able to synthesize many of them. **Tryptophan** is one such amino acid that *E. coli* can either ingest from the environment or synthesize using enzymes that are encoded by five genes. These five genes are next to each other in what is called the **tryptophan (*trp*) operon** ([\[link\]](#)). The genes are transcribed into a single mRNA, which is then translated to produce all five enzymes. If tryptophan is present in the environment, then *E. coli* does not need to synthesize it and the *trp* operon is switched off. However, when tryptophan availability is low, the switch controlling the operon is turned on, the mRNA is transcribed, the enzyme proteins are

translated, and tryptophan is synthesized.



The *trp* operon includes three important regions: the coding region, the *trp* operator and the *trp* promoter. The coding region includes the genes for the five tryptophan biosynthesis enzymes. Just before the coding region is the **transcriptional start site**. The promoter sequence, to which RNA polymerase binds to initiate transcription, is before or “upstream” of the transcriptional start site. Between the promoter and the transcriptional start site is the operator region.

The *trp* **operator** contains the DNA code to which the *trp* repressor protein can bind. However, the repressor alone cannot bind to the operator. When tryptophan is present in the cell, two tryptophan molecules bind to the *trp* repressor, which changes the shape of the repressor protein to a form that can bind to the *trp* operator. Binding of the tryptophan–

repressor complex at the operator physically prevents the RNA polymerase from binding to the promoter and transcribing the downstream genes.

When tryptophan is not present in the cell, the repressor by itself does not bind to the operator, the polymerase can transcribe the enzyme genes, and tryptophan is synthesized. Because the repressor protein actively binds to the operator to keep the genes turned off, the *trp* operon is said to be *negatively regulated* and the proteins that bind to the operator to silence *trp* expression are **negative regulators**.

Link to Learning

Watch this video to learn more about the *trp* operon.

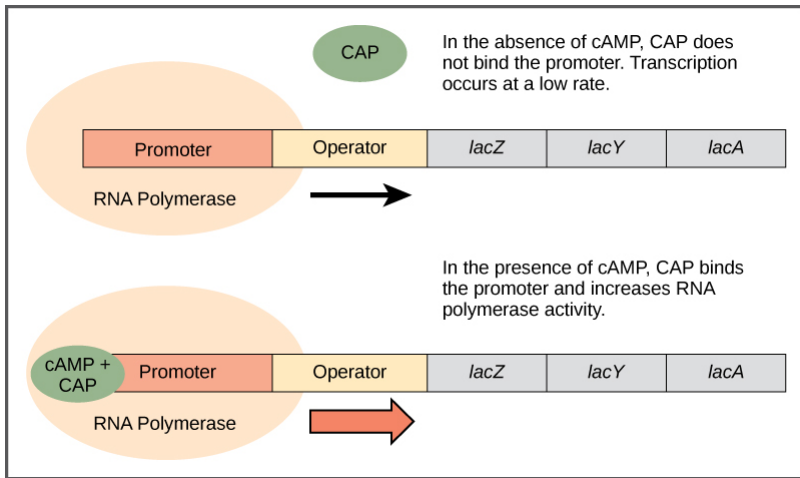
https://www.openstax.org/l/trp_operon

Transcriptional activation by the CAP protein. When glucose levels fall, *E. coli* may use other sugars for fuel but must transcribe new genes to do so. As glucose supplies become limited, cAMP levels increase. This cAMP binds to the CAP protein, a positive regulator that binds to a promoter region upstream of the genes required to use other sugar sources.

Catabolite Activator Protein (CAP): A Transcriptional Activator

Just as the *trp* operon is negatively regulated by tryptophan molecules, there are proteins that bind to the promoter sequences that act as **positive regulators** to turn genes on and activate them. For example, when glucose is scarce, *E. coli* bacteria can turn to other sugar sources for fuel. To do this, new genes to process these alternate sugars must be transcribed. When glucose levels drop, cyclic AMP (cAMP) begins to accumulate in the cell. The cAMP molecule is a signaling molecule that is involved in glucose and energy metabolism in *E. coli*.

Accumulating cAMP binds to the positive regulator **catabolite activator protein (CAP)**, a protein that binds to the promoters of operons which control the processing of alternative sugars. When cAMP binds to CAP, the complex then binds to the promoter region of the genes that are needed to use the alternate sugar sources ([\[link\]](#)). In these operons, a CAP-binding site is located upstream of the RNA-polymerase-binding site in the promoter. CAP binding stabilizes the binding of RNA polymerase to the promoter region and increases transcription of the associated protein-coding genes.



The *lac* Operon: An Inducible Operon

The third type of gene regulation in prokaryotic cells occurs through *inducible operons*, which have proteins that bind to activate or repress transcription depending on the local environment and the needs of the cell. The *lac* operon is a typical inducible operon. As mentioned previously, *E. coli* is able to use other sugars as energy sources when glucose concentrations are low. One such sugar source is lactose. The ***lac* operon** encodes the genes necessary to acquire and process the lactose from the local environment. The Z gene of the *lac* operon encodes beta-galactosidase, which breaks lactose down to glucose and galactose.

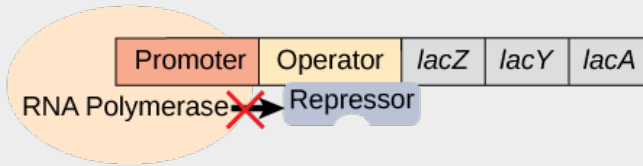
However, for the *lac* operon to be activated, two conditions must be met. First, the level of glucose

must be very low or non-existent. Second, lactose must be present. Only when glucose is absent and lactose is present will the *lac* operon be transcribed ([link]). In the absence of glucose, the binding of the CAP protein makes transcription of the *lac* operon more effective. When lactose is present, it binds to the *lac* repressor and changes its shape so that it cannot bind to the *lac* operator to prevent transcription. This combination of conditions makes sense for the cell, because it would be energetically wasteful to synthesize the enzymes to process lactose if glucose was plentiful or lactose was not available.

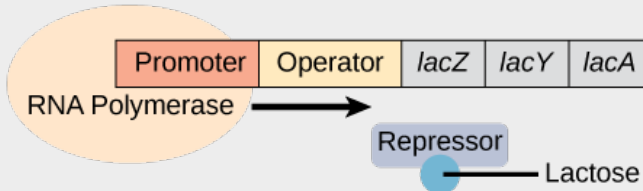
Visual Connection

Regulation of the *lac* operon. Transcription of the *lac* operon is carefully regulated so that its expression only occurs when glucose is limited and lactose is present to serve as an alternative fuel source.

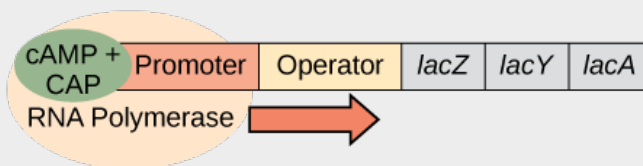
In the absence of lactose, the lac repressor binds the operator, and transcription is blocked.



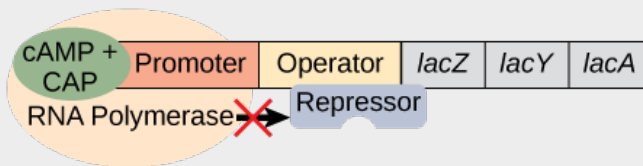
In the presence of lactose, the lac repressor is released from the operator, and transcription proceeds at a slow rate.



cAMP-CAP complex stimulates RNA Polymerase activity and increases RNA synthesis.



However, even in the presence of cAMP-CAP complex, RNA synthesis is blocked when repressor is bound to the operator.



In *E. coli*, the *trp* operon is on by default, while the *lac* operon is off. Why do you think this is the case?

If glucose is present, then CAP fails to bind to the promoter sequence to activate transcription. If lactose is absent, then the repressor binds to the operator to prevent transcription. If either of these conditions is met, then transcription remains off. Only when glucose is absent and lactose is present is the *lac* operon transcribed ([link](#)).

**Signals
that
Induce or
Repress
Transcription
of the *lac*
Operon**

Glucose	CAP binds	Lactose	Repressor binds	Transcription
+			+	No
+		+		Some
	+		+	No
-	+	+	-	Yes

Link to Learning
 Watch an animated tutorial about the workings of *lac* operon here.

Section Summary

The regulation of gene expression in prokaryotic cells occurs at the transcriptional level. There are two major kinds of proteins that control prokaryotic transcription: repressors and activators. Repressors bind to an operator region to block the action of RNA polymerase. Activators bind to the promoter to enhance the binding of RNA polymerase. Inducer molecules can increase transcription either by inactivating repressors or by activating activator proteins. In the *trp* operon, the *trp* repressor is itself activated by binding to tryptophan. Therefore, if tryptophan is not needed, the repressor is bound to the operator and transcription remains off. The *lac* operon is activated by the CAP (catabolite activator protein), which binds to the promoter to stabilize RNA polymerase binding. CAP is itself activated by cAMP, whose concentration rises as the concentration of glucose falls. However, the *lac* operon also requires the presence of lactose for transcription to occur. Lactose inactivates the *lac* repressor, and prevents the repressor protein from binding to the *lac* operator. With the repressor

inactivated, transcription may proceed. Therefore glucose must be absent and lactose must be present for effective transcription of the *lac* operon.

Visual Connection Questions

[\[link\]](#) In *E. coli*, the *trp* operon is on by default, while the *lac* operon is off. Why do you think that this is the case?

[\[link\]](#) Tryptophan is an amino acid essential for making proteins, so the cell always needs to have some on hand. However, if plenty of tryptophan is present, it is wasteful to make more, and the expression of the *trp* receptor is repressed. Lactose, a sugar found in milk, is not always available. It makes no sense to make the enzymes necessary to digest an energy source that is not available, so the *lac* operon is only turned on when lactose is present.

Review Questions

If glucose is absent, but so is lactose, the *lac*

operon will be _____.

1. activated
2. repressed
3. activated, but only partially
4. mutated

B

Prokaryotic cells lack a nucleus. Therefore, the genes in prokaryotic cells are:

1. all expressed, all of the time
2. transcribed and translated almost simultaneously
3. transcriptionally controlled because translation begins before transcription ends
4. b and c are both true

D

The *ara* operon is an inducible operon that controls the production of the sugar arabinose. When arabinose is present in a bacterium it binds to the protein AraC, and the complex binds to the initiator site to promote transcription. In this scenario, AraC is a(n) _____.

1. activator
 2. inducer
 3. repressor
 4. operator
-

A

Critical Thinking Questions

Describe how transcription in prokaryotic cells can be altered by external stimulation such as excess lactose in the environment.

Environmental stimuli can increase or induce transcription in prokaryotic cells. In this example, lactose in the environment will induce the transcription of the *lac* operon, but only if glucose is not available in the environment.

What is the difference between a repressible and an inducible operon?

A repressible operon uses a protein bound to the promoter region of a gene to keep the gene repressed or silent. This repressor must be

actively removed in order to transcribe the gene. An inducible operon is either activated or repressed depending on the needs of the cell and what is available in the local environment.

Glossary

activator

protein that binds to prokaryotic operators to increase transcription

catabolite activator protein (CAP)

protein that complexes with cAMP to bind to the promoter sequences of operons which control sugar processing when glucose is not available

inducible operon

operon that can be activated or repressed depending on cellular needs and the surrounding environment

lac operon

operon in prokaryotic cells that encodes genes required for processing and intake of lactose

negative regulator

protein that prevents transcription

operator

region of DNA outside of the promoter region

that binds activators or repressors that control gene expression in prokaryotic cells

operon

collection of genes involved in a pathway that are transcribed together as a single mRNA in prokaryotic cells

positive regulator

protein that increases transcription

repressor

protein that binds to the operator of prokaryotic genes to prevent transcription

transcriptional start site

site at which transcription begins

trp operon

series of genes necessary to synthesize tryptophan in prokaryotic cells

tryptophan

amino acid that can be synthesized by prokaryotic cells when necessary

Eukaryotic Epigenetic Gene Regulation

By the end of this section, you will be able to do the following:

- Explain how chromatin remodeling controls transcriptional access
- Describe how access to DNA is controlled by histone modification
- Describe how DNA methylation is related to epigenetic gene changes

Eukaryotic gene expression is more complex than prokaryotic gene expression because the processes of transcription and translation are physically separated. Unlike prokaryotic cells, eukaryotic cells can regulate gene expression at many different levels. Epigenetic changes are inheritable changes in gene expression that do not result from changes in the DNA sequence. Eukaryotic gene expression begins with control of access to the DNA.

Transcriptional access to the DNA can be controlled in two general ways: chromatin remodeling and DNA methylation. Chromatin remodeling changes the way that DNA is associated with chromosomal histones. DNA methylation is associated with developmental changes and gene silencing.

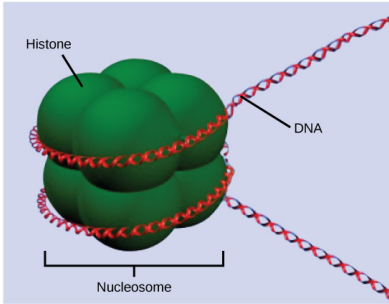
DNA is folded around histone proteins to create (a) nucleosome complexes. These nucleosomes control the access of proteins to the underlying DNA. When viewed through an electron microscope (b), the nucleosomes look like beads on a string. (credit

“micrograph”: modification of work by Chris Woodcock) Histone proteins and DNA nucleotides can be modified chemically. Modifications affect nucleosome spacing and gene expression. (credit: modification of work by NIH)

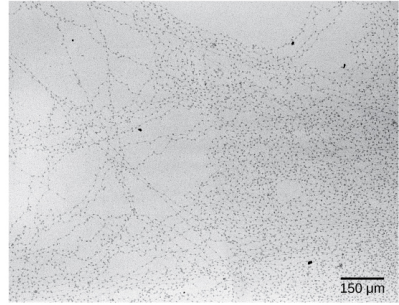
Epigenetic Control: Regulating Access to Genes within the Chromosome

The human genome encodes over 20,000 genes, with hundreds to thousands of genes on each of the 23 human chromosomes. The DNA in the nucleus is precisely wound, folded, and compacted into chromosomes so that it will fit into the nucleus. It is also organized so that specific segments can be accessed as needed by a specific cell type.

The first level of organization, or **packing**, is the winding of DNA strands around histone proteins. Histones package and order DNA into structural units called nucleosome complexes, which can control the access of proteins to the DNA regions ([\[link\]](#)a). Under the electron microscope, this winding of DNA around histone proteins to form nucleosomes looks like small beads on a string ([\[link\]](#)b).



(a)

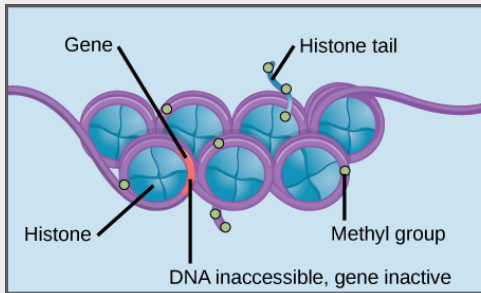


(b)

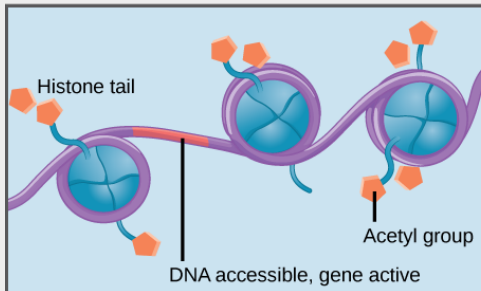
These beads (histone proteins) can move along the string (DNA) to expose different sections of the molecule. If DNA encoding a specific gene is to be transcribed into RNA, the nucleosomes surrounding that region of DNA can slide down the DNA to open that specific chromosomal region and allow for the transcriptional machinery (RNA polymerase) to initiate transcription ([\[link\]](#)).

Visual Connection

Nucleosomes can slide along DNA. When nucleosomes are spaced closely together (top), transcription factors cannot bind and gene expression is turned off. When the nucleosomes are spaced far apart (bottom), the DNA is exposed. Transcription factors can bind, allowing gene expression to occur. Modifications to the histones and DNA affect nucleosome spacing.



Methylation of DNA and histones causes nucleosomes to pack tightly together. Transcription factors cannot bind the DNA, and genes are not expressed.



Histone acetylation results in loose packing of nucleosomes. Transcription factors can bind the DNA and genes are expressed.

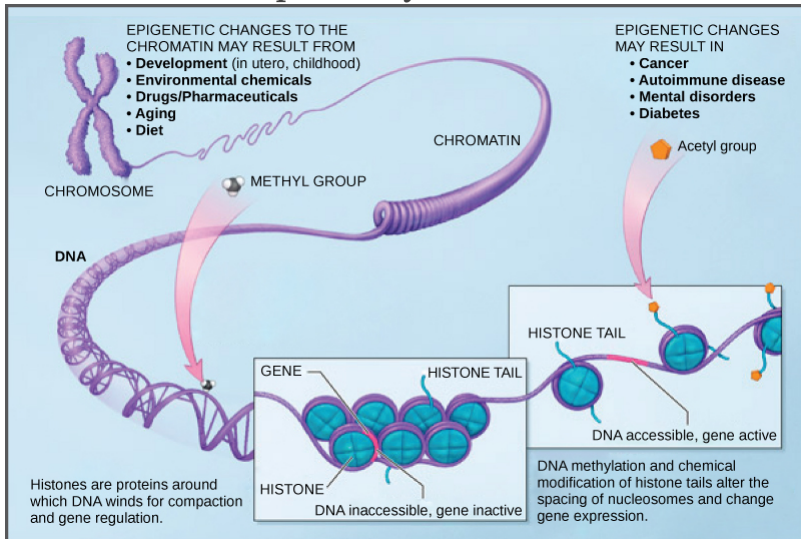
In females, one of the two X chromosomes is inactivated during embryonic development because of epigenetic changes to the chromatin. What impact do you think these changes would have on nucleosome packing?

How closely the histone proteins associate with the DNA is regulated by signals found on both the histone proteins and on the DNA. These signals are functional groups added to histone proteins or to DNA and determine whether a chromosomal region should be open or closed ([link](#) depicts modifications to histone proteins and DNA). These tags are not permanent, but may be added or removed as needed. Some chemical groups (phosphate, methyl, or acetyl groups) are attached

to specific amino acids in histone "tails" at the N-terminus of the protein. These groups do not alter the DNA base sequence, but they do alter how tightly wound the DNA is around the histone proteins. DNA is a negatively charged molecule and unmodified histones are positively charged; therefore, changes in the charge of the histone will change how tightly wound the DNA molecule will be. By adding chemical modifications like acetyl groups, the charge becomes less positive, and the binding of DNA to the histones is relaxed. Altering the location of nucleosomes and the tightness of histone binding opens some regions of chromatin to transcription and closes others.

The DNA molecule itself can also be modified by methylation. DNA methylation occurs within very specific regions called CpG islands. These are stretches with a high frequency of cytosine and guanine dinucleotide DNA pairs (CG) found in the promoter regions of genes. The cytosine member of the CG pair can be methylated (a methyl group is added). Methylated genes are usually silenced, although methylation may have other regulatory effects. In some cases, genes that are silenced during the development of the gametes of one parent are transmitted in their silenced condition to the offspring. Such genes are said to be imprinted. Parental diet or other environmental conditions may also affect the methylation patterns of genes, which in turn modifies gene expression. Changes in

chromatin organization interact with DNA methylation. DNA methyltransferases appear to be attracted to chromatin regions with specific histone modifications. Highly methylated (*hypermethylated*) DNA regions with deacetylated histones are tightly coiled and transcriptionally inactive.



Epigenetic changes are not permanent, although they often persist through multiple rounds of cell division and may even cross generational lines. Chromatin remodeling alters the chromosomal structure (open or closed) as needed. If a gene is to be transcribed, the histone proteins and DNA in the chromosomal region encoding that gene are modified in a way that opens the promoter region to allow RNA polymerase and other proteins, called **transcription factors**, to bind and initiate transcription. If a gene is to remain turned off, or silenced, the histone proteins and DNA have

different modifications that signal a closed chromosomal configuration. In this closed configuration, the RNA polymerase and transcription factors do not have access to the DNA and transcription cannot occur ([link]).

Link to Learning

View this video that describes how epigenetic regulation controls gene expression.

https://www.openstax.org/l/epigenetic_reg

Section Summary

In eukaryotic cells, the first stage of gene-expression control occurs at the epigenetic level. Epigenetic mechanisms control access to the chromosomal region to allow genes to be turned on or off.

Chromatin remodeling controls how DNA is packed into the nucleus by regulating how tightly the DNA is wound around histone proteins. The DNA itself may be methylated to selectively silence genes. The addition or removal of chemical modifications (or flags) to histone proteins or DNA signals the cell to open or close a chromosomal region. Therefore, eukaryotic cells can control whether a gene is

expressed by controlling accessibility to the binding of RNA polymerase and its transcription factors.

Visual Connection Questions

[\[link\]](#) In females, one of the two X chromosomes is inactivated during embryonic development because of epigenetic changes to the chromatin. What impact do you think these changes would have on nucleosome packing?

[\[link\]](#) The nucleosomes would pack more tightly together.

Review Questions

What are epigenetic modifications?

1. the addition of reversible changes to histone proteins and DNA
2. the removal of nucleosomes from the DNA
3. the addition of more nucleosomes to the DNA
4. mutation of the DNA sequence

A

Which of the following are true of epigenetic changes?

1. allow DNA to be transcribed
2. move histones to open or close a chromosomal region
3. are temporary
4. all of the above

D

Critical Thinking Questions

In cancer cells, alteration to epigenetic modifications turns off genes that are normally expressed. Hypothetically, how could you reverse this process to turn these genes back on?

You can create medications that reverse the epigenetic processes (to add histone acetylation marks or to remove DNA methylation) and create an open chromosomal configuration.

A scientific study demonstrated that rat mothering behavior impacts the stress response in their pups. Rats that were born and grew up with attentive mothers showed low activation of stress-response genes later in life, while rats with inattentive mothers had high activation of stress-response genes in the same situation. An additional study that swapped the pups at birth (i.e., rats born to inattentive mothers grew up with attentive mothers and vice versa) showed the same positive effect of attentive mothering. How do genetics and/or epigenetics explain the results of this study?

Swapping the pups at birth indicates that the genes inherited from the attentive or inattentive mothers do not explain the rats' stress-responses later in life. Instead, researchers found that the attentive mothering caused the methylation of genes that control the expression of stress receptors in the brain. Thus, rats that received attentive maternal care exhibited epigenetic changes that limited the expression of stress-response genes, and that the effect was durable over their lifespans.

Some autoimmune diseases show a positive correlation with dramatically decreased expression of histone deacetylase 9 (HDAC9, an enzyme that removes acetyl groups from

histones). Why would the decreased expression of HDAC9 cause immune cells to produce inflammatory genes at inappropriate times?

Histone acetylation reduces the positive charge of histone proteins, loosening the DNA wrapped around the histones. This looser DNA can then interact with transcription factors to express genes found in that region. Normally, once the gene is no longer needed, histone deacetylase enzymes remove the acetyl groups from histones so that the DNA becomes tightly wound and inaccessible again. However, when there is a defect in HDAC9, the deacetylation may not occur. In an immune cell, this would mean that inflammatory genes that were made accessible during an infection are not tightly rewound around the histones.

Glossary

transcription factor

protein that binds to the DNA at the promoter or enhancer region and that influences transcription of a gene

Eukaryotic Transcription Gene Regulation

By the end of this section, you will be able to do the following:

- Discuss the role of transcription factors in gene regulation
- Explain how enhancers and repressors regulate gene expression

Like prokaryotic cells, the transcription of genes in eukaryotes requires the action of an RNA polymerase to bind to a DNA sequence upstream of a gene in order to initiate transcription. However, unlike prokaryotic cells, the eukaryotic RNA polymerase requires other proteins, or transcription factors, to facilitate transcription initiation. RNA polymerase by itself cannot initiate transcription in eukaryotic cells. There are two types of transcription factors that regulate eukaryotic transcription:

General (or basal) transcription factors bind to the core promoter region to assist with the binding of RNA polymerase. *Specific transcription factors* bind to various regions outside of the core promoter region and interact with the proteins at the core promoter to enhance or repress the activity of the polymerase.

Link to Learning

View the process of transcription—the making of RNA from a DNA template.

The Promoter and the Transcription Machinery

Genes are organized to make the control of gene expression easier. The promoter region is immediately upstream of the coding sequence. This region can be short (only a few nucleotides in length) or quite long (hundreds of nucleotides long). The longer the promoter, the more available space for proteins to bind. This also adds more control to the transcription process. The length of the promoter is gene-specific and can differ dramatically between genes. Consequently, the level of control of gene expression can also differ quite dramatically between genes. The purpose of the **promoter** is to bind transcription factors that control the initiation of transcription.

Within the core promoter region, 25 to 35 bases upstream of the transcriptional start site, resides the TATA box. The TATA box has the consensus sequence of 5'-TATAAA-3'. The TATA box is the binding site for a protein complex called TFIID, which contains a TATA-binding protein. Binding of TFIID recruits other transcription factors, including

TFIIB, TFIIE, TFIIF, and TFIIH. Some of these transcription factors help to bind the RNA polymerase to the promoter, and others help to activate the transcription initiation complex.

In addition to the TATA box, other binding sites are found in some promoters. Some biologists prefer to restrict the range of the eukaryotic promoter to the core promoter, or polymerase binding site, and refer to these additional sites as promoter-proximal elements, because they are usually found within a few hundred base pairs upstream of the transcriptional start site. Examples of these elements are the CAAT box, with the consensus sequence 5'-CCAAT-3' and the GC box, with the consensus sequence 5'-GGGCGG-3'. Specific transcription factors can bind to these promoter-proximal elements to regulate gene transcription. A given gene may have its own combination of these specific transcription-factor binding sites. There are hundreds of transcription factors in a cell, each of which binds specifically to a particular DNA sequence motif. When transcription factors bind to the promoter just upstream of the encoded gene, it is referred to as a **cis-acting element**, because it is on the same chromosome just next to the gene. Transcription factors respond to environmental stimuli that cause the proteins to find their binding sites and initiate transcription of the gene that is needed.

Interaction between proteins at the promoter and

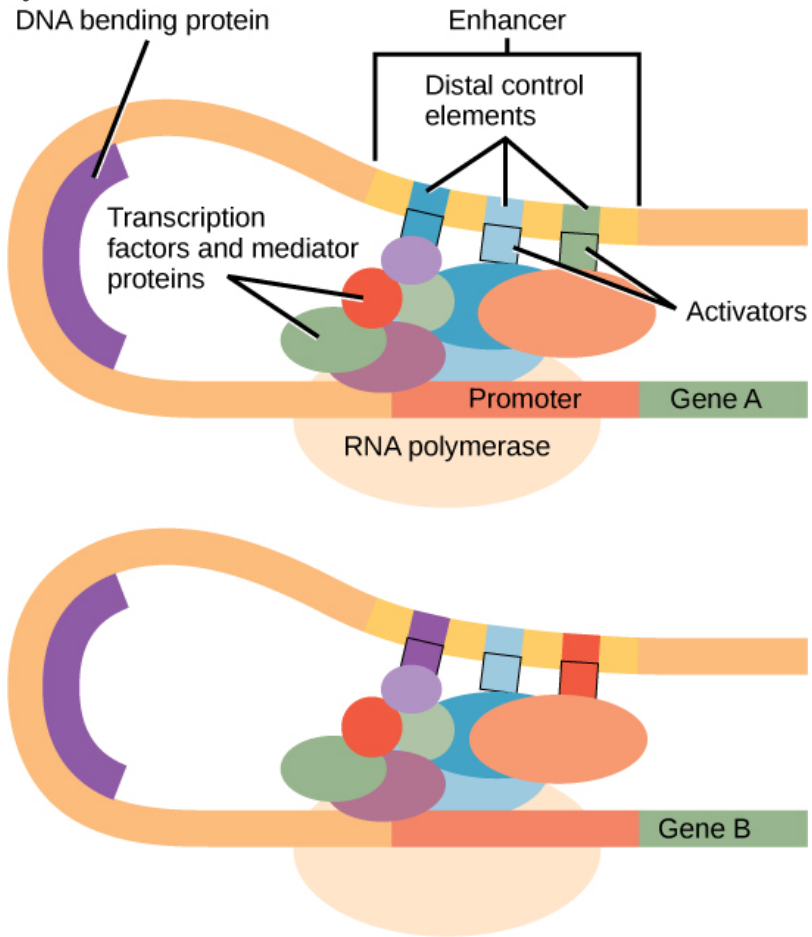
enhancer sites. An enhancer is a DNA sequence that promotes transcription. Each enhancer is made up of short DNA sequences called distal control elements. Activators bound to the distal control elements interact with mediator proteins and transcription factors. Two different genes may have the same promoter but different distal control elements, enabling differential gene expression.

Enhancers and Transcription

In some eukaryotic genes, there are additional regions that help increase or enhance transcription. These regions, called **enhancers**, are not necessarily close to the genes they enhance. They can be located upstream of a gene, within the coding region of the gene, downstream of a gene, or may be thousands of nucleotides away.

Enhancer regions are binding sequences, or sites, for specific transcription factors. When a protein transcription factor binds to its enhancer sequence, the shape of the protein changes, allowing it to interact with proteins at the promoter site. However, since the enhancer region may be distant from the promoter, the DNA must bend to allow the proteins at the two sites to come into contact. DNA bending proteins help to bend the DNA and bring the enhancer and promoter regions together ([link]). This shape change allows for the interaction of the specific activator proteins bound

to the enhancers with the general transcription factors bound to the promoter region and the RNA polymerase.



Turning Genes Off: Transcriptional Repressors

Like prokaryotic cells, eukaryotic cells also have mechanisms to prevent transcription.

Transcriptional repressors can bind to promoter or enhancer regions and block transcription. Like the transcriptional activators, repressors respond to external stimuli to prevent the binding of activating transcription factors.

Section Summary

To start transcription, general transcription factors, such as TFIID, TFIIB, and others, must first bind to the TATA box and recruit RNA polymerase to that location. Additional transcription factors may also bind to other regulatory elements at the promoter to increase or prevent transcription. In addition to promoter sequences, enhancer regions help augment transcription. Enhancers can be upstream, downstream, within a gene itself, or on other chromosomes. Specific transcription factors bound to enhancer regions may either increase or prevent transcription.

Review Questions

The binding of _____ is required for transcription to start.

1. a protein

2. DNA polymerase
 3. RNA polymerase
 4. a transcription factor
-

C

What will result from the binding of a transcription factor to an enhancer region?

1. decreased transcription of an adjacent gene
 2. increased transcription of a distant gene
 3. alteration of the translation of an adjacent gene
 4. initiation of the recruitment of RNA polymerase
-

B

A scientist compares the promoter regions of two genes. Gene A's core promoter plus proximal promoter elements encompasses 70bp. Gene B's core promoter plus proximal promoter elements encompasses 250bp. Which of the scientist's hypotheses is most likely to be correct?

1. More transcripts will be made from Gene B.

2. Transcription of Gene A involves fewer transcription factors.
 3. Enhancers control Gene B's transcription.
 4. Transcription of Gene A is more controlled than transcription of Gene B.
-

B

Critical Thinking Questions

A mutation within the promoter region can alter transcription of a gene. Describe how this can happen.

A mutation in the promoter region can change the binding site for a transcription factor that normally binds to increase transcription. The mutation could either decrease the ability of the transcription factor to bind, thereby decreasing transcription, or it can increase the ability of the transcription factor to bind, thus increasing transcription.

What could happen if a cell had too much of an activating transcription factor present?

If too much of an activating transcription factor were present, then transcription would be increased in the cell. This could lead to dramatic alterations in cell function.

A scientist identifies a potential transcription regulation site 300bp downstream of a gene and hypothesizes that it is a repressor. What experiment (with results) could he perform to support this hypothesis?

The easiest way to test his hypothesis would be to mutate the site in a cell, and monitor levels of the mRNA transcript made from the gene. If the levels of transcript increase in the mutated cell, then the site was repressing transcription.

Glossary

cis-acting element

transcription factor binding sites within the promoter that regulate the transcription of a gene adjacent to it

enhancer

segment of DNA that is upstream, downstream, perhaps thousands of nucleotides away, or on another chromosome

that influence the transcription of a specific gene

trans-acting element

transcription factor binding site found outside the promoter or on another chromosome that influences the transcription of a particular gene

transcription factor binding site

sequence of DNA to which a transcription factor binds

Eukaryotic Post-transcriptional Gene Regulation

By the end of this section, you will be able to do the following:

- Understand RNA splicing and explain its role in regulating gene expression
- Describe the importance of RNA stability in gene regulation

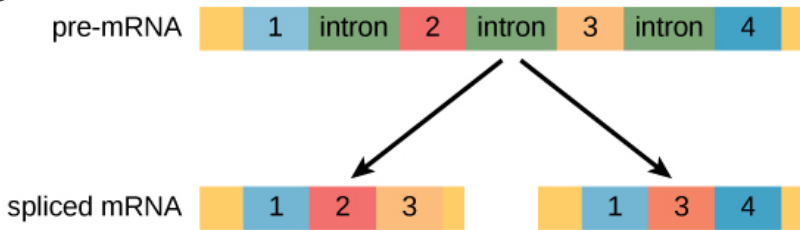
RNA is transcribed, but must be processed into a mature form before translation can begin. This processing that takes place after an RNA molecule has been transcribed, but before it is translated into a protein, is called *post-transcriptional modification*. As with the epigenetic and transcriptional stages of processing, this post-transcriptional step can also be regulated to control gene expression in the cell. If the RNA is not processed, shuttled, or translated, then no protein will be synthesized.

Pre-mRNA can be alternatively spliced to create different proteins.

RNA Splicing, the First Stage of Post-transcriptional Control

In eukaryotic cells, the RNA transcript often contains regions, called introns, that are removed prior to translation. The regions of RNA that code for protein are called **exons**. ([\[link\]](#)). After an RNA molecule has been transcribed, but prior to its

departure from the nucleus to be translated, the RNA is processed and the introns are removed by splicing. Splicing is done by spliceosomes, ribonucleoprotein complexes that can recognize the two ends of the intron, cut the transcript at those two points, and bring the exons together for ligation.



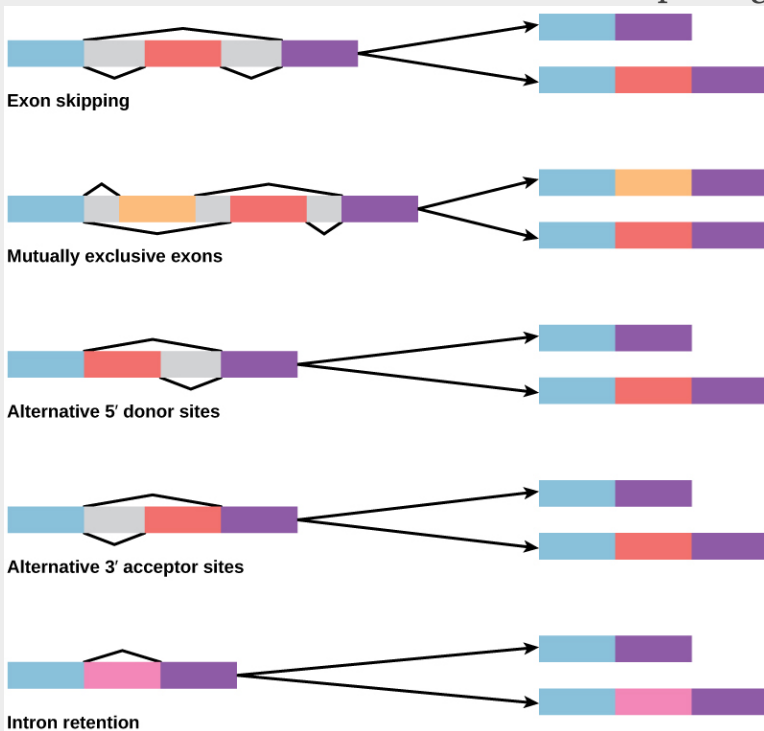
Evolution Connection

Alternative RNA Splicing

In the 1970s, genes were first observed that exhibited alternative RNA splicing. Alternative RNA splicing is a mechanism that allows different protein products to be produced from one gene when different combinations of exons are combined to form the mRNA ([\[link\]](#)). This alternative splicing can be haphazard, but more often it is controlled and acts as a mechanism of *gene regulation*, with the frequency of different splicing alternatives controlled by the cell as a way to control the production of different protein products in different cells or at different stages of development. Alternative splicing is now understood to be a common mechanism of gene

regulation in eukaryotes; according to one estimate, 70 percent of genes in humans are expressed as multiple proteins through alternative splicing. Although there are multiple ways to alternatively splice RNA transcripts, the original 5'-3' order of the exons is *always conserved*. That is, a transcript with exons 1 2 3 4 5 6 7 might be spliced 1 2 4 5 6 7 or 1 2 3 6 7, but never 1 2 5 4 3 6 7.

There are five basic modes of alternative splicing.



How could alternative splicing evolve? Introns have a beginning- and ending-recognition sequence; it is easy to imagine the failure of the splicing mechanism to identify the end of an intron and instead find the end of the next intron, thus

removing two introns and the intervening exon. In fact, there are mechanisms in place to prevent such intron skipping, but mutations are likely to lead to their failure. Such “mistakes” would more than likely produce a nonfunctional protein. Indeed, the cause of many genetic diseases is abnormal splicing rather than mutations in a coding sequence.

However, alternative splicing could possibly create a protein variant without the loss of the original protein, opening up possibilities for adaptation of the new variant to new functions. Gene duplication has played an important role in the evolution of new functions in a similar way by providing genes that may evolve without eliminating the original, functional protein.

Question: In the corn snake *Pantherophis guttatus*, there are several different color variants, including amelanistic snakes whose skin patterns display only red and yellow pigments. The cause of amelanism in these snakes was recently identified as the insertion of a transposable element into an intron in the OCA2 (oculocutaneous albinism) gene. How might the insertion of extra genetic material into an intron lead to a nonfunctional protein?

Link to Learning

Visualize how mRNA splicing happens by watching the process in action in this video.

https://www.openstax.org/l/mRNA_splicing

RNA-binding proteins. The protein-coding region of this processed mRNA is flanked by 5' and 3' untranslated regions (UTRs). The presence of RNA-binding proteins at the 5' or 3' UTR influences the stability of the RNA molecule.

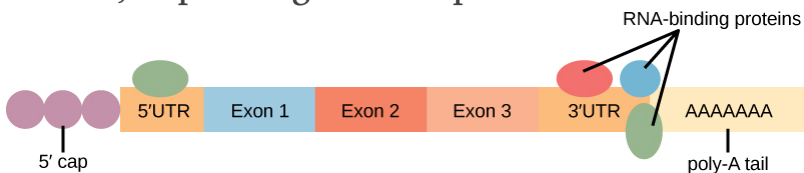
Control of RNA Stability

Before the mRNA leaves the nucleus, it is given two protective "caps" that prevent the ends of the strand from degrading during its journey. 5' and 3' exonucleases can degrade unprotected RNAs. The **5' cap**, which is placed on the 5' end of the mRNA, is usually composed of a methylated guanosine triphosphate molecule (GTP). The GTP is placed "backward" on the 5' end of the mRNA, so that the 5' carbons of the GTP and the terminal nucleotide are linked through three phosphates. The **poly-A tail**, which is attached to the 3' end, is usually composed of a long chain of adenine nucleotides. These changes protect the two ends of the RNA from exonuclease attack.

Once the RNA is transported to the cytoplasm, the length of time that the RNA resides there can be controlled. Each RNA molecule has a defined lifespan and decays at a specific rate. This rate of decay can influence how much protein is in the cell. If the decay rate is increased, the RNA will not exist in the cytoplasm as long, shortening the time available for translation of the mRNA to occur.

Conversely, if the rate of decay is decreased, the mRNA molecule will reside in the cytoplasm longer and more protein can be translated. This rate of decay is referred to as the RNA stability. If the RNA is stable, it will be detected for longer periods of time in the cytoplasm.

Binding of proteins to the RNA can also influence its stability. Proteins called **RNA-binding proteins**, or RBPs, can bind to the regions of the mRNA just upstream or downstream of the protein-coding region. These regions in the RNA that are not translated into protein are called the **untranslated regions**, or UTRs. They are not introns (those have been removed in the nucleus). Rather, these are regions that regulate mRNA localization, stability, and protein translation. The region just before the protein-coding region is called the **5' UTR**, whereas the region after the coding region is called the **3' UTR** ([\[link\]](#)). The binding of RBPs to these regions can increase or decrease the stability of an RNA molecule, depending on the specific RBP that binds.



RNA Stability and microRNAs

In addition to RBPs that bind to and control (increase or decrease) RNA stability, other elements

called microRNAs can bind to the RNA molecule. These **microRNAs**, or miRNAs, are short RNA molecules that are only 21 to 24 nucleotides in length. The miRNAs are made in the nucleus as longer pre-miRNAs. These pre-miRNAs are chopped into mature miRNAs by a protein called **Dicer**. Like transcription factors and RBPs, mature miRNAs recognize a specific sequence and bind to the RNA; however, miRNAs also associate with a ribonucleoprotein complex called the **RNA-induced silencing complex (RISC)**. The RNA component of the RISC base-pairs with complementary sequences on an mRNA and either impede translation of the message or lead to the degradation of the mRNA.

Section Summary

Post-transcriptional control can occur at any stage after transcription, including RNA splicing and RNA stability. Once RNA is transcribed, it must be processed to create a mature RNA that is ready to be translated. This involves the removal of introns that do not code for protein. Spliceosomes bind to the signals that mark the exon/intron border to remove the introns and ligate the exons together. Once this occurs, the RNA is mature and can be translated. Alternative splicing can produce more than one mRNA from a given transcript. Different splicing variants may be produced under different conditions.

RNA is created and spliced in the nucleus, but needs to be transported to the cytoplasm to be translated. RNA is transported to the cytoplasm through the nuclear pore complex. Once the RNA is in the cytoplasm, the length of time it resides there before being degraded, called RNA stability, can also be altered to control the overall amount of protein that is synthesized. The RNA stability can be increased, leading to longer residency time in the cytoplasm, or decreased, leading to shortened time and less protein synthesis. RNA stability is controlled by RNA-binding proteins (RPBs) and microRNAs (miRNAs). These RPBs and miRNAs bind to the 5' UTR or the 3' UTR of the RNA to increase or decrease RNA stability. MicroRNAs associated with RISC complexes may repress translation or lead to mRNA breakdown.

Review Questions

Which of the following are involved in post-transcriptional control?

1. control of RNA splicing
 2. control of RNA shuttling
 3. control of RNA stability
 4. all of the above
-

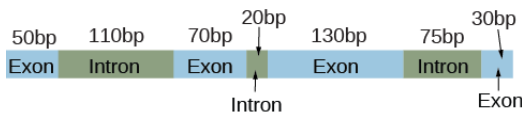
D

Binding of an RNA binding protein will _____ the stability of the RNA molecule.

1. increase
2. decrease
3. neither increase nor decrease
4. either increase or decrease

D

An unprocessed pre-mRNA has the following structure.



Which of the following is not a possible size (in bp) of the mature mRNA?

1. 205bp
2. 180bp
3. 150bp
4. 100bp

A

Alternative splicing has been estimated to occur in more than 95% of multi-exon genes. Which of the following is not an evolutionary advantage of alternative splicing?

1. Alternative splicing increases diversity without increasing genome size.
2. Different gene isoforms can be expressed in different tissues.
3. Alternative splicing creates shorter mRNA transcripts.
4. Different gene isoforms can be expressed during different stages of development.

C

Critical Thinking Questions

Describe how RBPs can prevent miRNAs from degrading an RNA molecule.

RNA binding proteins (RBP) bind to the RNA and can either increase or decrease the stability of the RNA. If they increase the stability of the RNA molecule, the RNA will remain intact in

the cell for a longer period of time than normal. Since both RBPs and miRNAs bind to the RNA molecule, RBP can potentially bind first to the RNA and prevent the binding of the miRNA that will degrade it.

How can external stimuli alter post-transcriptional control of gene expression?

External stimuli can modify RNA-binding proteins (i.e., through phosphorylation of proteins) to alter their activity.

Glossary

3' UTR

3' untranslated region; region just downstream of the protein-coding region in an RNA molecule that is not translated

5' cap

a methylated guanosine triphosphate (GTP) molecule that is attached to the 5' end of a messenger RNA to protect the end from degradation

5' UTR

5' untranslated region; region just upstream of the protein-coding region in an RNA molecule

that is not translated

Dicer

enzyme that chops the pre-miRNA into the mature form of the miRNA

microRNA (miRNA)

small RNA molecules (approximately 21 nucleotides in length) that bind to RNA molecules to degrade them

poly-A tail

a series of adenine nucleotides that are attached to the 3' end of an mRNA to protect the end from degradation

RNA-binding protein (RBP)

protein that binds to the 3' or 5' UTR to increase or decrease the RNA stability

RNA stability

how long an RNA molecule will remain intact in the cytoplasm

untranslated region

segment of the RNA molecule that is not translated into protein. These regions lie before (upstream or 5') and after (downstream or 3') the protein-coding region

RISC

protein complex that binds along with the

miRNA to the RNA to degrade it

Eukaryotic Translational and Post-translational Gene Regulation

By the end of this section, you will be able to do the following:

- Understand the process of translation and discuss its key factors
- Describe how the initiation complex controls translation
- Explain the different ways in which the post-translational control of gene expression takes place

After RNA has been transported to the cytoplasm, it is translated into protein. Control of this process is largely dependent on the RNA molecule. As previously discussed, the stability of the RNA will have a large impact on its translation into a protein. As the stability changes, the amount of time that it is available for translation also changes.

The Initiation Complex and Translation Rate

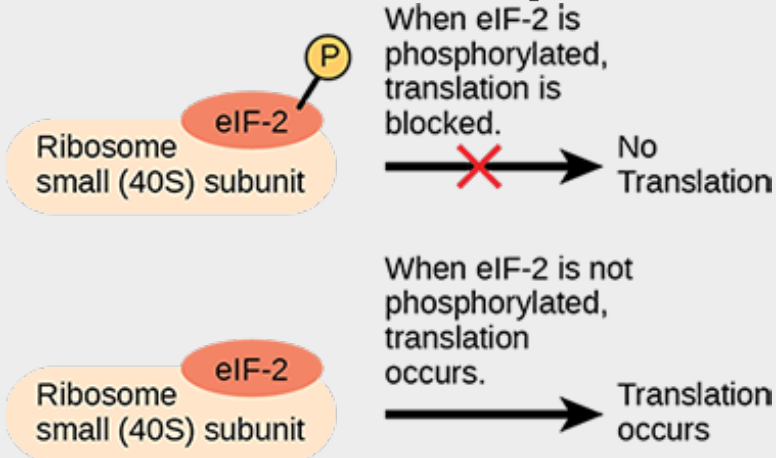
Like transcription, translation is controlled by proteins that bind and initiate the process. In translation, the complex that assembles to start the process is referred to as the translation **initiation complex**. In eukaryotes, translation is initiated by

binding the initiating met-tRNA_i to the 40S ribosome. This tRNA is brought to the 40S ribosome by a protein initiation factor, **eukaryotic initiation factor-2 (eIF-2)**. The eIF-2 protein binds to the high-energy molecule **guanosine triphosphate (GTP)**. The tRNA-eIF2-GTP complex then binds to the 40S ribosome. A second complex forms on the mRNA. Several different initiation factors recognize the 5' cap of the mRNA and proteins bound to the poly-A tail of the same mRNA, forming the mRNA into a loop. The cap-binding protein eIF4F brings the mRNA complex together with the 40S ribosome complex. The ribosome then scans along the mRNA until it finds a start codon AUG. When the anticodon of the initiator tRNA and the start codon are aligned, the GTP is hydrolyzed, the initiation factors are released, and the large **60S ribosomal subunit** binds to form the translation complex. The binding of eIF-2 to the RNA is controlled by phosphorylation. If eIF-2 is phosphorylated, it undergoes a conformational change and cannot bind to GTP. Therefore, the initiation complex cannot form properly and translation is impeded ([\[link\]](#)). When eIF-2 remains unphosphorylated, the initiation complex can form normally and translation can proceed.

Visual Connection

Gene expression can be controlled by factors that

bind the translation initiation complex.



An increase in phosphorylation levels of eIF-2 has been observed in patients with neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's. What impact do you think this might have on protein synthesis?

Proteins with ubiquitin tags are marked for degradation within the proteasome.

Chemical Modifications, Protein Activity, and Longevity

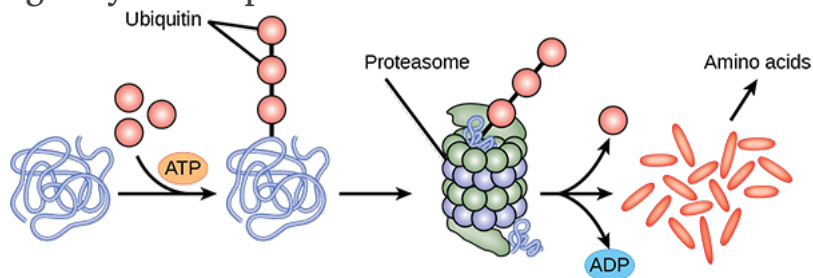
Proteins can be chemically modified with the addition of groups including methyl, phosphate, acetyl, and ubiquitin groups. The addition or removal of these groups from proteins regulates their activity or the length of time they exist in the

cell. Sometimes these modifications can regulate where a protein is found in the cell—for example, in the nucleus, in the cytoplasm, or attached to the plasma membrane.

Chemical modifications occur in response to external stimuli such as stress, the lack of nutrients, heat, or ultraviolet light exposure. These changes can alter epigenetic accessibility, transcription, mRNA stability, or translation—all resulting in changes in expression of various genes. This is an efficient way for the cell to rapidly change the levels of specific proteins in response to the environment. Because proteins are involved in every stage of gene regulation, the phosphorylation of a protein (depending on the protein that is modified) can alter accessibility to the chromosome, can alter translation (by altering transcription factor binding or function), can change nuclear shuttling (by influencing modifications to the nuclear pore complex), can alter RNA stability (by binding or not binding to the RNA to regulate its stability), can modify translation (increase or decrease), or can change post-translational modifications (add or remove phosphates or other chemical modifications).

The addition of an ubiquitin group to a protein marks that protein for degradation. Ubiquitin acts like a flag indicating that the protein lifespan is complete. These proteins are moved to the

proteasome, an organelle that functions to remove proteins, to be degraded ([\[link\]](#)). One way to control gene expression, therefore, is to alter the longevity of the protein.



Section Summary

Changing the status of the RNA or the protein itself can affect the amount of protein, the function of the protein, or how long it is found in the cell. To translate the protein, a protein initiator complex must assemble on the RNA. Modifications (such as phosphorylation) of proteins in this complex can prevent proper translation from occurring. Once a protein has been synthesized, it can be modified (phosphorylated, acetylated, methylated, or ubiquitinated). These post-translational modifications can greatly impact the stability, degradation, or function of the protein.

Visual Connection Questions

[\[link\]](#) An increase in phosphorylation levels of eIF-2 has been observed in patients with neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's. What impact do you think this might have on protein synthesis?

[\[link\]](#) Protein synthesis would be inhibited.

Review Questions

Post-translational modifications of proteins can affect which of the following?

1. protein function
2. transcriptional regulation
3. chromatin modification
4. all of the above

A

A scientist mutates eIF-2 to eliminate its GTP hydrolysis capability. How would this mutated form of eIF-2 alter translation?

1. Initiation factors would not be able to bind

- to mRNA.
2. The large ribosomal subunit would not be able to interact with mRNA transcripts.
 3. tRNAⁱ-Met would not scan mRNA transcripts for the start codon.
 4. eIF-2 would not be able to interact with the small ribosomal subunit.
-

B

Critical Thinking Questions

Protein modification can alter gene expression in many ways. Describe how phosphorylation of proteins can alter gene expression.

Because proteins are involved in every stage of gene regulation, phosphorylation of a protein (depending on the protein that is modified) can alter accessibility to the chromosome, can alter translation (by altering the transcription factor binding or function), can change nuclear shuttling (by influencing modifications to the nuclear pore complex), can alter RNA stability (by binding or not binding to the RNA to regulate its stability), can modify translation

(increase or decrease), or can change post-translational modifications (add or remove phosphates or other chemical modifications).

Alternative forms of a protein can be beneficial or harmful to a cell. What do you think would happen if too much of an alternative protein bound to the 3' UTR of an RNA and caused it to degrade?

If the RNA degraded, then less of the protein that the RNA encodes would be translated. This could have dramatic implications for the cell.

Changes in epigenetic modifications alter the accessibility and transcription of DNA. Describe how environmental stimuli, such as ultraviolet light exposure, could modify gene expression.

Environmental stimuli, like ultraviolet light exposure, can alter the modifications to the histone proteins or DNA. Such stimuli may change an actively transcribed gene into a silenced gene by removing acetyl groups from histone proteins or by adding methyl groups to DNA.

A scientist discovers a virus encoding a Protein X that degrades a subunit of the eIF4F complex. Knowing that this virus transcribes its own mRNAs in the cytoplasm of human cells, why would Protein X be an effective virulence factor?

Degrading the eIF4F complex prevents the pre-initiation complex (eIF-2-GTP, tRNAⁱ-Met, and 40S ribosomal subunit) from being recruited to the 5' cap of mature mRNAs in the cell. This allows the virus to hijack the translation machinery of the human cell to translate its own (uncapped) mRNA transcripts instead.

Glossary

eukaryotic initiation factor-2 (eIF-2)

protein that binds first to an mRNA to initiate translation

guanine diphosphate (GDP)

molecule that is left after the energy is used to start translation

guanine triphosphate (GTP)

energy-providing molecule that binds to eIF-2 and is needed for translation

initiation complex

protein complex containing eIF-2 that starts translation

large 60S ribosomal subunit

second, larger ribosomal subunit that binds to the RNA to translate it into protein

proteasome

organelle that degrades proteins

small 40S ribosomal subunit

ribosomal subunit that binds to the RNA to translate it into protein

Cancer and Gene Regulation

By the end of this section, you will be able to do the following:

- Describe how changes to gene expression can cause cancer
- Explain how changes to gene expression at different levels can disrupt the cell cycle
- Discuss how understanding regulation of gene expression can lead to better drug design

Cancer is not a single disease but includes many different diseases. In cancer cells, mutations modify cell-cycle control and cells don't stop growing as they normally would. Mutations can also alter the growth rate or the progression of the cell through the cell cycle. One example of a gene modification that alters the growth rate is increased phosphorylation of cyclin B, a protein that controls the progression of a cell through the cell cycle and serves as a cell-cycle checkpoint protein.

For cells to move through each phase of the cell cycle, the cell must pass through checkpoints. This ensures that the cell has properly completed the step and has not encountered any mutation that will alter its function. Many proteins, including cyclin B, control these checkpoints. The phosphorylation of cyclin B, a post-translational event, alters its function. As a result, cells can progress through the cell cycle unimpeded, even if mutations exist in the

cell and its growth should be terminated. This post-translational change of cyclin B prevents it from controlling the cell cycle and contributes to the development of cancer.

Cancer: Disease of Altered Gene Expression

Cancer can be described as a disease of altered gene expression. There are many proteins that are turned on or off (gene activation or gene silencing) that dramatically alter the overall activity of the cell. A gene that is not normally expressed in that cell can be switched on and expressed at high levels. This can be the result of gene mutation or changes in gene regulation (epigenetic, transcription, post-transcription, translation, or post-translation).

Changes in epigenetic regulation, transcription, RNA stability, protein translation, and post-translational control can be detected in cancer. While these changes don't occur simultaneously in one cancer, changes at each of these levels can be detected when observing cancer at different sites in different individuals. Therefore, changes in **histone acetylation** (epigenetic modification that leads to gene silencing), activation of transcription factors by phosphorylation, increased RNA stability, increased translational control, and protein modification can all be detected at some point in various cancer cells.

Scientists are working to understand the common changes that give rise to certain types of cancer or how a modification might be exploited to destroy a tumor cell.

Tumor Suppressor Genes, Oncogenes, and Cancer

In normal cells, some genes function to prevent excess, inappropriate cell growth. These are tumor-suppressor genes, which are active in normal cells to prevent uncontrolled cell growth. There are many tumor-suppressor genes in cells. The most studied tumor-suppressor gene is p53, which is mutated in over 50 percent of all cancer types. The p53 protein itself functions as a transcription factor. It can bind to sites in the promoters of genes to initiate transcription. Therefore, the mutation of p53 in cancer will dramatically alter the transcriptional activity of its target genes.

Link to Learning

Watch [this animation](#) to learn more about the use of p53 in fighting cancer.

Proto-oncogenes are positive cell-cycle regulators. When mutated, proto-oncogenes can become

oncogenes and cause cancer. Overexpression of the oncogene can lead to uncontrolled cell growth. This is because oncogenes can alter transcriptional activity, stability, or protein translation of another gene that directly or indirectly controls cell growth. An example of an oncogene involved in cancer is a protein called myc. **Myc** is a transcription factor that is aberrantly activated in Burkett's Lymphoma, a cancer of the lymph system. Overexpression of myc transforms normal B cells into cancerous cells that continue to grow uncontrollably. High B-cell numbers can result in tumors that can interfere with normal bodily function. Patients with Burkett's lymphoma can develop tumors on their jaw or in their mouth that interfere with the ability to eat.

Cancer and Epigenetic Alterations

Silencing genes through epigenetic mechanisms is also very common in cancer cells. There are characteristic modifications to histone proteins and DNA that are associated with silenced genes. In cancer cells, the DNA in the promoter region of silenced genes is methylated on cytosine DNA residues in CpG islands. Histone proteins that surround that region lack the acetylation modification that is present when the genes are expressed in normal cells. This combination of DNA methylation and histone deacetylation (epigenetic modifications that lead to gene silencing) is

commonly found in cancer. When these modifications occur, the gene present in that chromosomal region is silenced. Increasingly, scientists understand how epigenetic changes are altered in cancer. Because these changes are temporary and can be reversed—for example, by preventing the action of the histone deacetylase protein that removes acetyl groups, or by DNA methyl transferase enzymes that add methyl groups to cytosines in DNA—it is possible to design new drugs and new therapies to take advantage of the reversible nature of these processes. Indeed, many researchers are testing how a silenced gene can be switched back on in a cancer cell to help re-establish normal growth patterns.

Genes involved in the development of many other illnesses, ranging from allergies to inflammation to autism, are thought to be regulated by epigenetic mechanisms. As our knowledge of how genes are controlled deepens, new ways to treat diseases like cancer will emerge.

Cancer and Transcriptional Control

Alterations in cells that give rise to cancer can affect the transcriptional control of gene expression. Mutations that activate transcription factors, such as increased phosphorylation, can increase the binding of a transcription factor to its binding site in a

promoter. This could lead to increased transcriptional activation of that gene that results in modified cell growth. Alternatively, a mutation in the DNA of a promoter or enhancer region can increase the binding ability of a transcription factor. This could also lead to the increased transcription and aberrant gene expression that is seen in cancer cells.

Researchers have been investigating how to control the transcriptional activation of gene expression in cancer. Identifying how a transcription factor binds, or a pathway that activates where a gene can be turned off, has led to new drugs and new ways to treat cancer. In breast cancer, for example, many proteins are overexpressed. This can lead to increased phosphorylation of key transcription factors that increase transcription. One such example is the overexpression of the epidermal growth-factor receptor (EGFR) in a subset of breast cancers. The EGFR pathway activates many protein kinases that, in turn, activate many transcription factors which control genes involved in cell growth. New drugs that prevent the activation of EGFR have been developed and are used to treat these cancers.

Cancer and Post-transcriptional Control

Changes in the post-transcriptional control of a gene can also result in cancer. Recently, several groups of

researchers have shown that specific cancers have altered expression of miRNAs. Because miRNAs bind to the 3' UTR of RNA molecules to degrade them, overexpression of these miRNAs could be detrimental to normal cellular activity. Too many miRNAs could dramatically decrease the RNA population, leading to a decrease in protein expression. Several studies have demonstrated a change in the miRNA population in specific cancer types. It appears that the subset of miRNAs expressed in breast cancer cells is quite different from the subset expressed in lung cancer cells or even from normal breast cells. This suggests that alterations in miRNA activity can contribute to the growth of breast cancer cells. These types of studies also suggest that if some miRNAs are specifically expressed only in cancer cells, they could be potential drug targets. It would, therefore, be conceivable that new drugs that turn off miRNA expression in cancer could be an effective method to treat cancer.

Cancer and Translational/Post-translational Control

There are many examples of how translational or post-translational modifications of proteins arise in cancer. Modifications are found in cancer cells from the increased translation of a protein to changes in protein phosphorylation to alternative splice

variants of a protein. An example of how the expression of an alternative form of a protein can have dramatically different outcomes is seen in colon cancer cells. The c-Flip protein, a protein involved in mediating the cell-death pathway, comes in two forms: long (c-FLIPL) and short (c-FLIPS). Both forms appear to be involved in initiating controlled cell-death mechanisms in normal cells. However, in colon cancer cells, expression of the long form results in increased cell growth instead of cell death. Clearly, the expression of the wrong protein dramatically alters cell function and contributes to the development of cancer.

New Drugs to Combat Cancer: Targeted Therapies

Scientists are using what is known about the regulation of gene expression in disease states, including cancer, to develop new ways to treat and prevent disease development. Many scientists are designing drugs on the basis of the gene expression patterns within individual tumors. This idea, that therapy and medicines can be tailored to an individual, has given rise to the field of personalized medicine. With an increased understanding of gene regulation and gene function, medicines can be designed to specifically target diseased cells without harming healthy cells. Some new medicines, called

targeted therapies, have exploited the overexpression of a specific protein or the mutation of a gene to develop a new medication to treat disease. One such example is the use of anti-EGF receptor medications to treat the subset of breast cancer tumors that have very high levels of the EGF protein. Undoubtedly, more targeted therapies will be developed as scientists learn more about how gene expression changes can cause cancer.

Career Connection

Clinical Trial Coordinator

A clinical trial coordinator is the person managing the proceedings of the clinical trial. This job includes coordinating patient schedules and appointments, maintaining detailed notes, building the database to track patients (especially for long-term follow-up studies), ensuring proper documentation has been acquired and accepted, and working with the nurses and doctors to facilitate the trial and publication of the results. A clinical trial coordinator may have a science background, like a nursing degree, or other certification. People who have worked in science labs or in clinical offices are also qualified to become a clinical trial coordinator. These jobs are generally in hospitals; however, some clinics and doctor's offices also conduct clinical trials and may hire a coordinator.

Section Summary

Cancer can be described as a disease of altered gene expression. Changes at every level of eukaryotic gene expression can be detected in some form of cancer at some point in time. In order to understand how changes to gene expression can cause cancer, it is critical to understand how each stage of gene regulation works in normal cells. By understanding the mechanisms of control in normal, non-diseased cells, it will be easier for scientists to understand what goes wrong in disease states including complex ones like cancer.

Review Questions

Cancer causing genes are called _____.

1. transformation genes
2. tumor suppressor genes
3. oncogenes
4. mutated genes

Targeted therapies are used in patients with a set gene expression pattern. A targeted therapy that prevents the activation of the estrogen receptor in breast cancer would be beneficial to which type of patient?

1. patients who express the EGFR receptor in normal cells
2. patients with a mutation that inactivates the estrogen receptor
3. patients with lots of the estrogen receptor expressed in their tumor
4. patients that have no estrogen receptor expressed in their tumor

C

Critical Thinking Questions

New drugs are being developed that decrease DNA methylation and prevent the removal of acetyl groups from histone proteins. Explain how these drugs could affect gene expression to help kill tumor cells.

These drugs will keep the histone proteins and

the DNA methylation patterns in the open chromosomal configuration so that transcription is feasible. If a gene is silenced, these drugs could reverse the epigenetic configuration to re-express the gene.

How can understanding the gene expression pattern in a cancer cell tell you something about that specific form of cancer?

Understanding which genes are expressed in a cancer cell can help diagnose the specific form of cancer. It can also help identify treatment options for that patient. For example, if a breast cancer tumor expresses the EGFR in high numbers, it might respond to specific anti-EGFR therapy. If that receptor is not expressed, it would not respond to that therapy.

Glossary

DNA methylation

epigenetic modification that leads to gene silencing; a process involving adding a methyl group to the DNA molecule

histone acetylation

epigenetic modification that leads to gene silencing; a process involving adding or

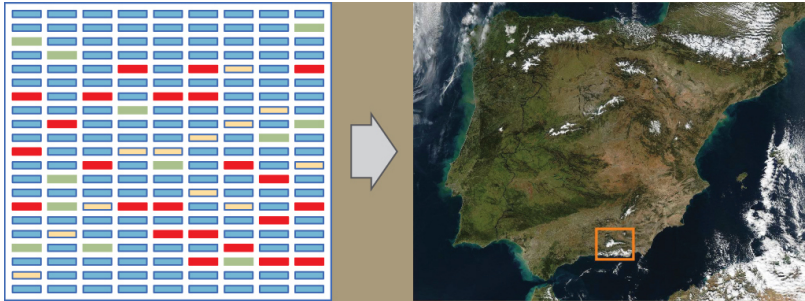
removing an acetyl functional group

myc

oncogene that causes cancer in many cancer cells

Introduction

class = "introduction" Genomics compares the DNA of different organisms, enabling scientists to create maps with which to navigate different organisms' DNA. (credit "map": modification of photo by NASA)



The study of nucleic acids began with the discovery of DNA, progressed to the study of genes and small fragments, and has now exploded to the field of genomics. Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. DNA sequencing technology has contributed to advances in genomics. Just as information technology has led to Google maps that enable people to obtain detailed information about locations around the globe, researchers use genomic information to create similar DNA maps of different organisms. These findings have helped anthropologists to better understand human migration and have aided the medical field through mapping human genetic diseases. Genomic information can contribute to scientific understanding in various ways and

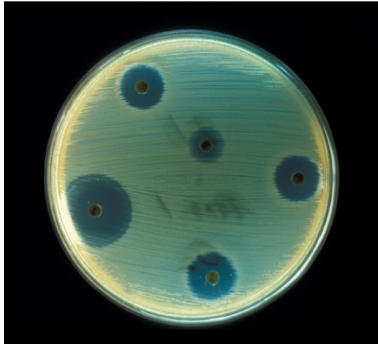
knowledge in the field is quickly growing.

Biotechnology

By the end of this section, you will be able to do the following:

- Describe gel electrophoresis
- Explain molecular and reproductive cloning
- Describe biotechnology uses in medicine and agriculture

Biotechnology is the use of biological agents for technological advancement. Biotechnology was used for breeding livestock and crops long before people understood the scientific basis of these techniques. Since the discovery of the structure of DNA in 1953, the biotechnology field has grown rapidly through both academic research and private companies. The primary applications of this technology are in medicine (vaccine and antibiotic production) and agriculture (crop genetic modification in order to increase yields). Biotechnology also has many industrial applications, such as fermentation, treating oil spills, and producing biofuels ([\[link\]](#)). Fungi, bacteria, and other organisms that have antimicrobial properties produce antibiotics. The first antibiotic discovered was penicillin. Pharmaceutical companies now commercially produce and test antibiotics for their potential to inhibit bacterial growth. (credit "advertisement": modification of work by NIH; credit "test plate": modification of work by Don Stalons/CDC; scale-bar data from Matt Russell)



This diagram shows the basic method of DNA extraction. a) Shown are DNA fragments from seven samples run on a gel, stained with a fluorescent dye, and viewed under UV light; and b) a researcher from International Rice Research Institute, reviewing DNA profiles using UV light. (credit: a: James Jacob, Tompkins Cortland Community College b: International Rice Research Institute)

Scientists use polymerase chain reaction, or PCR, to amplify a specific DNA sequence. Primers—short pieces of DNA complementary to each end of the target sequence combine with genomic DNA, Taq polymerase, and deoxynucleotides. Taq polymerase is a DNA polymerase isolated from the thermostable bacterium *Thermus aquaticus* that is able to withstand the high temperatures that scientists use in PCR. *Thermus aquaticus* grows in the Lower Geyser Basin of Yellowstone National Park. Reverse transcriptase PCR (RT-PCR) is similar to PCR, but cDNA is made from an RNA template before PCR begins. Scientists use Southern blotting to find a particular sequence in a DNA sample. Scientists separate DNA fragments on a gel, transfer them to a nylon membrane, and incubate them with a DNA

probe complementary to the sequence of interest. Northern blotting is similar to Southern blotting, but scientists run RNA on the gel instead of DNA. In Western blotting, scientists run proteins on a gel and detect them using antibodies.

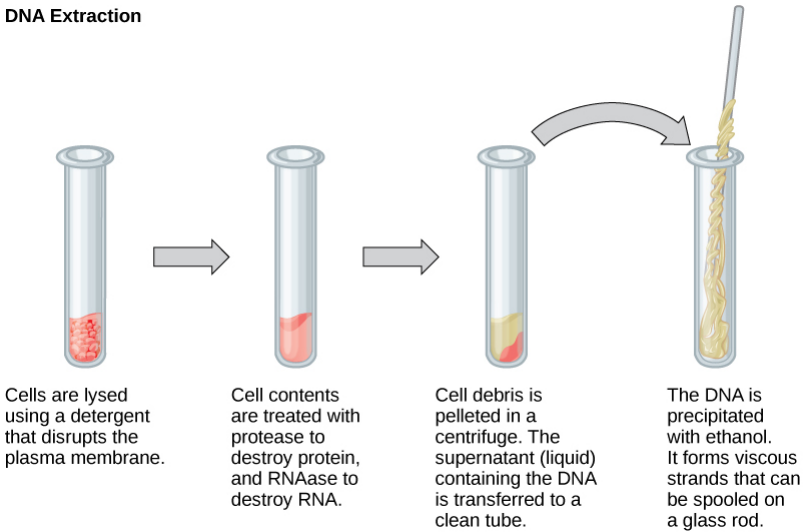
Basic Techniques to Manipulate Genetic Material (DNA and RNA)

To understand the basic techniques used to work with nucleic acids, remember that nucleic acids are macromolecules made of nucleotides (a sugar, a phosphate, and a nitrogenous base) linked by phosphodiester bonds. The phosphate groups on these molecules each have a net negative charge. An entire set of DNA molecules in the nucleus is called the genome. DNA has two complementary strands linked by hydrogen bonds between the paired bases. Exposure to high temperatures (DNA denaturation) can separate the two strands and cooling can reanneal them. The DNA polymerase enzyme can replicate the DNA. Unlike DNA, which is located in the eukaryotic cells' nucleus, RNA molecules leave the nucleus. The most common type of RNA that researchers analyze is the messenger RNA (mRNA) because it represents the protein-coding genes that are actively expressed. However, RNA molecules present some other challenges to analysis, as they are often less stable than DNA.

DNA and RNA Extraction

To study or manipulate nucleic acids, one must first isolate or extract the DNA or RNA from the cells. Researchers use various techniques to extract different types of DNA ([\[link\]](#)). Most nucleic acid extraction techniques involve steps to break open the cell and use enzymatic reactions to destroy all macromolecules that are not desired (such as unwanted molecule degradation and separation from the DNA sample). A **lysis buffer** (a solution which is mostly a detergent) breaks cells. Note that lysis means "to split". These enzymes break apart lipid molecules in the cell membranes and nuclear membranes. Enzymes such as **proteases** that break down proteins inactivate macromolecules, and **ribonucleases** (RNAses) that break down RNA. Using alcohol precipitates the DNA. Human genomic DNA is usually visible as a gelatinous, white mass. One can store the DNA samples frozen at -80°C for several years.

DNA Extraction

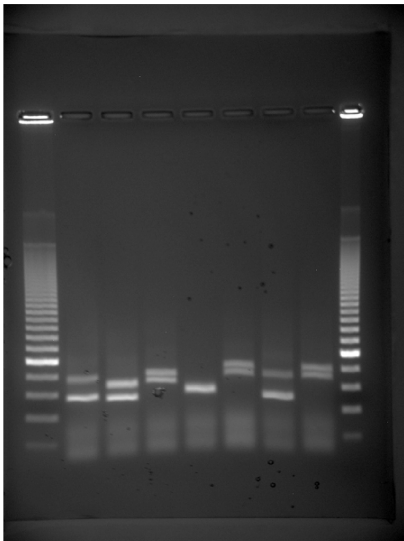


Scientists perform RNA analysis to study gene expression patterns in cells. RNA is naturally very unstable because RNases are commonly present in nature and very difficult to inactivate. Similar to DNA, RNA extraction involves using various buffers and enzymes to inactivate macromolecules and preserve the RNA.

Gel Electrophoresis

Because nucleic acids are negatively charged ions at neutral or basic pH in an aqueous environment, an electric field can mobilize them. **Gel electrophoresis** is a technique that scientists use to separate molecules on the basis of size, using this charge. One can separate the nucleic acids as whole chromosomes or fragments. The nucleic acids load into a slot near the semisolid, porous gel matrix's

negative electrode, and pulled toward the positive electrode at the gel's opposite end. Smaller molecules move through the gel's pores faster than larger molecules. This difference in the migration rate separates the fragments on the basis of size. There are molecular weight standard samples that researchers can run alongside the molecules to provide a size comparison. We can observe nucleic acids in a gel matrix using various fluorescent or colored dyes. Distinct nucleic acid fragments appear as bands at specific distances from the gel's top (the negative electrode end) on the basis of their size ([link]). A mixture of genomic DNA fragments of varying sizes appear as a long smear; whereas, uncut genomic DNA is usually too large to run through the gel and forms a single large band at the gel's top.



(a)



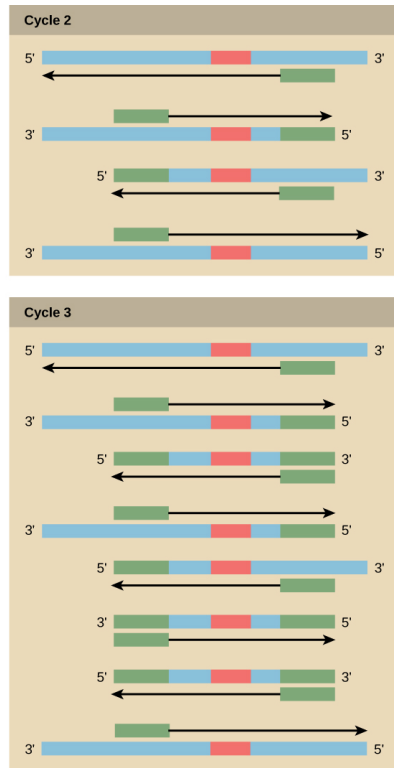
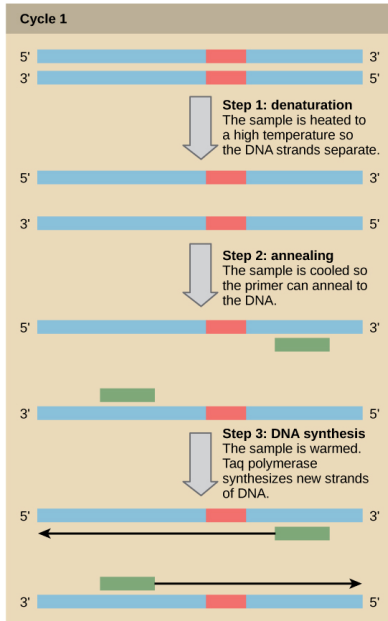
(b)

Nucleic Acid Fragment Amplification by Polymerase Chain Reaction

Although genomic DNA is visible to the naked eye when it is extracted in bulk, DNA analysis often requires focusing on one or more specific genome regions. **Polymerase chain reaction (PCR)** is a technique that scientists use to amplify specific DNA regions for further analysis ([\[link\]](#)). Researchers use PCR for many purposes in laboratories, such as cloning gene fragments to analyze genetic diseases, identifying contaminant foreign DNA in a sample, and amplifying DNA for sequencing. More practical applications include determining paternity and detecting genetic diseases.

Polymerase Chain Reaction (PCR)

The PCR cycle consists of three steps—denaturation, annealing, and DNA synthesis—that occur at high, low, and intermediate temperatures, respectively. The cycle is repeated again and again, resulting in a doubling of DNA molecules each time. After several cycles, the vast majority of strands produced are the same length as the distance between the two primers.



DNA fragments can also be amplified from an RNA template in a process called **reverse transcriptase PCR (RT-PCR)**. The first step is to recreate the original DNA template strand (called cDNA) by applying DNA nucleotides to the mRNA. This process is called reverse transcription. This requires the presence of an enzyme called reverse transcriptase. After the cDNA is made, regular PCR can be used to amplify it.

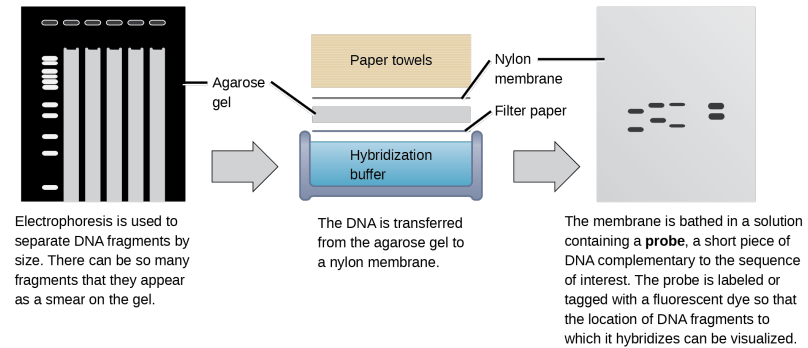
Link to Learning

Deepen your understanding of the polymerase chain reaction by clicking through [this interactive exercise](#).

Hybridization, Southern Blotting, and Northern Blotting

Scientists can probe nucleic acid samples, such as fragmented genomic DNA and RNA extracts, for the presence of certain sequences. Scientists design and label short DNA fragments, or **probes** with radioactive or fluorescent dyes to aid detection. Gel electrophoresis separates the nucleic acid fragments according to their size. Scientists then transfer the fragments in the gel onto a nylon membrane in a procedure we call **blotting** ([link](#)). Scientists can then probe the nucleic acid fragments that are bound to the membrane's surface with specific radioactively or fluorescently labeled probe sequences. When scientists transfer DNA to a nylon membrane, they refer to the technique as **Southern blotting**. When they transfer the RNA to a nylon membrane, they call it **Northern blotting**. Scientists use Southern blots to detect the presence of certain DNA sequences in a given genome, and Northern blots to detect gene expression.

Southern Blotting



Molecular Cloning

In general, the word “cloning” means the creation of a perfect replica; however, in biology, the re-creation of a whole organism is referred to as “reproductive cloning.” Long before attempts were made to clone an entire organism, researchers learned how to reproduce desired regions or fragments of the genome, a process that is referred to as molecular cloning.

Cloning small genome fragments allows researchers to manipulate and study specific genes (and their protein products), or noncoding regions in isolation. A plasmid, or vector, is a small circular DNA molecule that replicates independently of the chromosomal DNA. In cloning, scientists can use the plasmid molecules to provide a "folder" in which to insert a desired DNA fragment. Plasmids are usually introduced into a bacterial host for proliferation. In the bacterial context, scientists call the DNA

fragment from the human genome (or the genome of another studied organism) **foreign DNA**, or a transgene, to differentiate it from the bacterium's DNA, or the **host DNA**.

Plasmids occur naturally in bacterial populations (such as *Escherichia coli*) and have genes that can contribute favorable traits to the organism, such as **antibiotic resistance** (the ability to be unaffected by antibiotics). Scientists have repurposed and engineered plasmids as vectors for molecular cloning and the large-scale production of important reagents, such as insulin and human growth hormone. An important feature of plasmid vectors is the ease with which scientists can introduce a foreign DNA fragment via the **multiple cloning site (MCS)**. The MCS is a short DNA sequence containing multiple sites that different commonly available restriction endonucleases can cut. **Restriction endonucleases** recognize specific DNA sequences and cut them in a predictable manner. They are naturally produced by bacteria as a defense mechanism against foreign DNA. Many restriction endonucleases make staggered cuts in the two DNA strands, such that the cut ends have a 2- or 4-base single-stranded overhang. Because these overhangs are capable of annealing with complementary overhangs, we call them “sticky ends.” Adding the enzyme DNA ligase permanently joins the DNA fragments via phosphodiester bonds. In this way, scientists can splice any DNA fragment generated by

restriction endonuclease cleavage between the plasmid DNA's two ends that has been cut with the same restriction endonuclease ([\[link\]](#)).

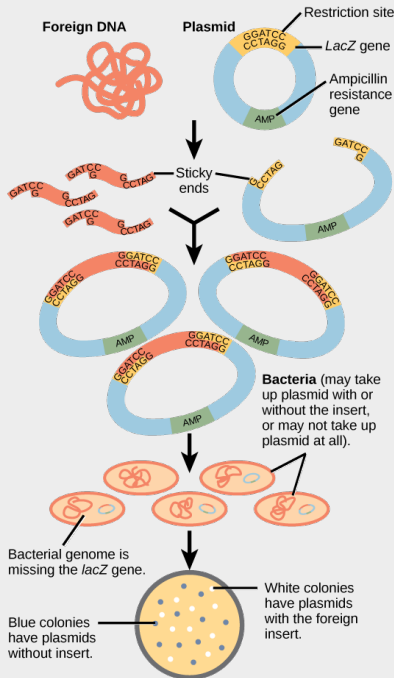
Recombinant DNA Molecules

Plasmids with foreign DNA inserted into them are called **recombinant DNA** molecules because they are created artificially and do not occur in nature. They are also called chimeric molecules because the origin of different molecule parts of the molecules can be traced back to different species of biological organisms or even to chemical synthesis. We call proteins that are expressed from recombinant DNA molecules **recombinant proteins**. Not all recombinant plasmids are capable of expressing genes. The recombinant DNA may need to move into a different vector (or host) that is better designed for gene expression. Scientists may also engineer plasmids to express proteins only when certain environmental factors stimulate them, so they can control the recombinant proteins' expression.

Visual Connection

This diagram shows the steps involved in molecular cloning.

Molecular Cloning



The foreign DNA and plasmid are cut with the same **restriction enzyme**, which recognizes a particular sequence of DNA called a **restriction site**. The restriction site occurs only once in the plasmid, and is located within the *lacZ* gene, a gene necessary for metabolizing lactose.

The restriction enzyme creates sticky ends that allow the foreign DNA and cloning vector to anneal. An enzyme called **ligase** glues the annealed fragments together.

The ligated cloning vector is transformed into a bacterial host strain that is ampicillin sensitive and is missing the *lacZ* gene from its genome.

Bacteria are grown on media containing ampicillin and X-gal, a chemical that is metabolized by the same pathway as lactose. The ampicillin kills bacteria without plasmid. Plasmids lacking the foreign insert have an intact *lacZ* gene and are able to metabolize X-gal, releasing a dye that turns the colony blue. Plasmids with an insert have a disrupted *lacZ* gene and produce white colonies.

You are working in a molecular biology lab and, unbeknownst to you, your lab partner left the foreign genomic DNA that you are planning to clone on the lab bench overnight instead of storing it in the freezer. As a result, it was degraded by nucleases, but still used in the experiment. The plasmid, on the other hand, is fine. What results would you expect from your molecular cloning experiment?

1. There will be no colonies on the bacterial plate.
2. There will be blue colonies only.
3. There will be blue and white colonies.
4. There will be white colonies only.

Link to Learning

View an [animation of recombination in cloning](#) from the DNA Learning Center.

Cellular Cloning

Unicellular organisms, such as bacteria and yeast, naturally produce clones of themselves when they replicate asexually by binary fission; this is known as **cellular cloning**. The nuclear DNA duplicates by the process of mitosis, which creates an exact replica of the genetic material.

Reproductive Cloning

Reproductive cloning is a method scientists use to clone or identically copy an entire multicellular organism. Most multicellular organisms undergo reproduction by sexual means, which involves genetic hybridization of two individuals (parents), making it impossible to generate an identical copy or a clone of either parent. Recent advances in biotechnology have made it possible to artificially induce mammal asexual reproduction in the laboratory.

Parthenogenesis, or “virgin birth,” occurs when an embryo grows and develops without egg fertilization. This is a form of asexual reproduction. An example of parthenogenesis occurs in species in which the female lays an egg and if the egg is fertilized, it is a diploid egg and the individual develops into a female. If the egg is not fertilized, it remains a haploid egg and develops into a male. The unfertilized egg is a parthenogenic, or virgin egg. Some insects and reptiles lay parthenogenic eggs that can develop into adults.

Sexual reproduction requires two cells. When the haploid egg and sperm cells fuse, a diploid zygote results. The zygote nucleus contains the genetic information to produce a new individual. However, early embryonic development requires the cytoplasmic material contained in the egg cell. This idea forms the basis for reproductive cloning. Therefore, if we replace the egg cell's haploid nucleus with a diploid nucleus from the cell of any individual of the same species (a donor), it will become a zygote that is genetically identical to the donor. Somatic cell nuclear transfer is the technique of transferring a diploid nucleus into an enucleated egg. Scientists can use it for either therapeutic cloning or reproductive cloning.

The first cloned animal was Dolly, a sheep born in 1996. The reproductive cloning success rate at the time was very low. Dolly lived for seven years and

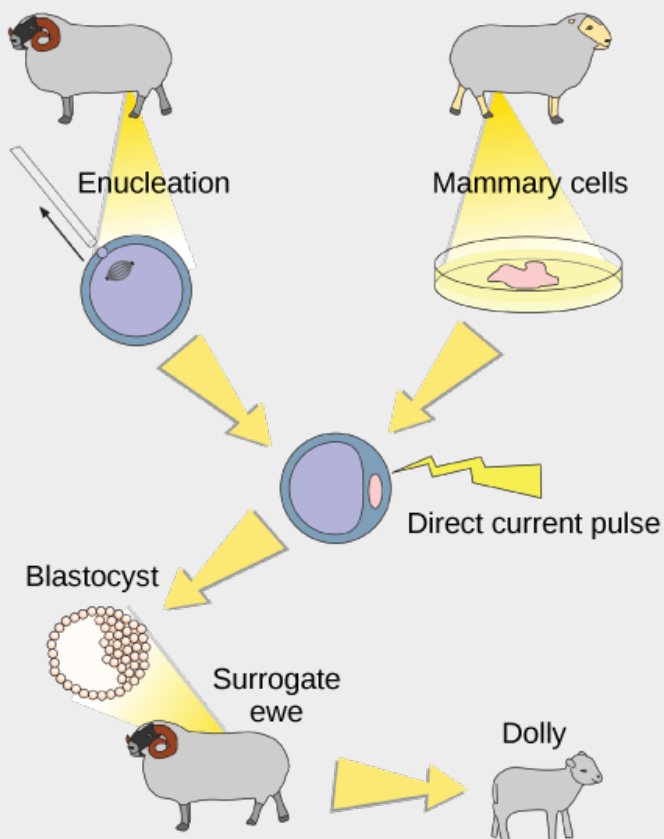
died of respiratory complications ([\[link\]](#)). There is speculation that because the cell DNA belongs to an older individual, DNA's age may affect a cloned individual's life expectancy. Since Dolly, scientists have cloned successfully several animals such as horses, bulls, and goats, although these animals often exhibit facial, limb, and cardiac abnormalities. There have been attempts at producing cloned human embryos as sources of embryonic stem cells for therapeutic purposes. Therapeutic cloning produces stem cells in the attempt to remedy detrimental diseases or defects (unlike reproductive cloning, which aims to reproduce an organism). Still, some have met therapeutic cloning efforts with resistance because of bioethical considerations.

Visual Connection

Dolly the sheep was the first mammal to be cloned. To create Dolly, they removed the nucleus from a donor egg cell. They then introduced the nucleus from a second sheep into the cell, which divided to the blastocyst stage before they implanted it in a surrogate mother. (credit: modification of work by "Squidonium"/Wikimedia Commons)

Scottish Blackface
(Cytoplasmic donor)

Finn-Dorset
(Nuclear donor)



Do you think Dolly was a Finn-Dorset or a Scottish Blackface sheep?

Genetic Engineering

Genetic engineering is the alteration of an organism's genotype using recombinant DNA technology to modify an organism's DNA to achieve desirable traits. The addition of foreign DNA in the form of recombinant DNA vectors generated by molecular cloning is the most common method of genetic engineering. The organism that receives the recombinant DNA is a **genetically modified organism** (GMO). If the foreign DNA comes from a different species, the host organism is **transgenic**. Scientists have genetically modified bacteria, plants, and animals since the early 1970s for academic, medical, agricultural, and industrial purposes. In the US, GMOs such as Roundup-ready soybeans and borer-resistant corn are part of many common processed foods.

Gene Targeting

Although classical methods of studying gene function began with a given phenotype and determined the genetic basis of that phenotype, modern techniques allow researchers to start at the DNA sequence level and ask: "What does this gene or DNA element do?" This technique, reverse genetics, has resulted in reversing the classic genetic methodology. This method would be similar to damaging a body part to determine its function. An insect that loses a wing cannot fly, which means that the wing's function is flight. The classical genetic method would compare insects that cannot

fly with insects that can fly, and observe that the non-flying insects have lost wings. Similarly, mutating or deleting genes provides researchers with clues about gene function. We collectively call the methods they use to disable gene function gene targeting. **Gene targeting** is the use of recombinant DNA vectors to alter a particular gene's expression, either by introducing mutations in a gene, or by eliminating a certain gene's expression by deleting a part or all of the gene sequence from the organism's genome.

Gene therapy using an adenovirus vector can be used to cure certain genetic diseases in which a person has a defective gene. (credit: NIH)

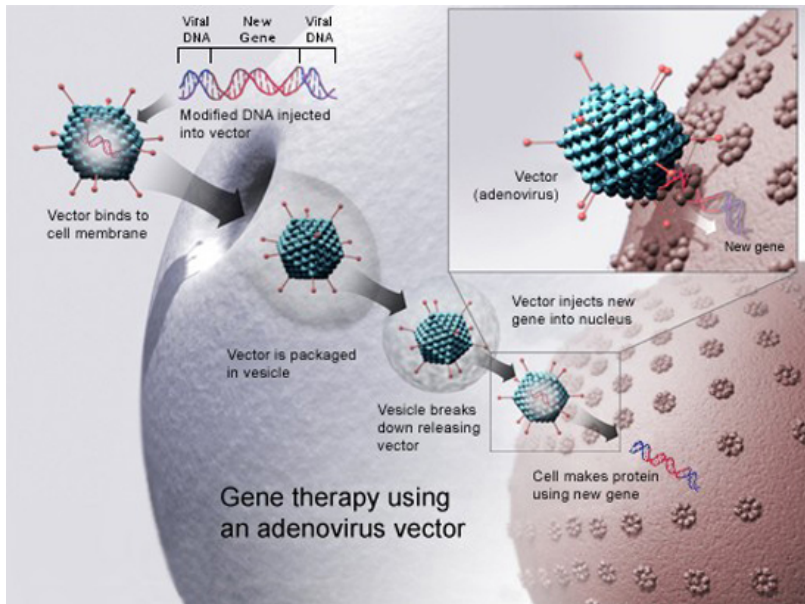
Biotechnology in Medicine and Agriculture

It is easy to see how biotechnology can be used for medicinal purposes. Knowledge of the genetic makeup of our species, the genetic basis of heritable diseases, and the invention of technology to manipulate and fix mutant genes provides methods to treat the disease. Biotechnology in agriculture can enhance resistance to disease, pest, and environmental stress, and improve both crop yield and quality.

Genetic Diagnosis and Gene Therapy

Scientists call the process of testing for suspected genetic defects before administering treatment **genetic diagnosis** by **genetic testing**. Depending on the inheritance patterns of a disease-causing gene, family members are advised to undergo genetic testing. For example, doctors usually advise women diagnosed with breast cancer to have a biopsy so that the medical team can determine the genetic basis of cancer development. Doctors base treatment plans on genetic test findings that determine the type of cancer. If inherited gene mutations cause the cancer, doctors also advise other female relatives to undergo genetic testing and periodic screening for breast cancer. Doctors also offer genetic testing for fetuses (or embryos with in vitro fertilization) to determine the presence or absence of disease-causing genes in families with specific debilitating diseases.

Gene therapy is a genetic engineering technique used to cure disease. In its simplest form, it involves the introduction of a good gene at a random location in the genome to aid the cure of a disease that is caused by a mutated gene. The good gene is usually introduced into diseased cells as part of a vector transmitted by a virus that can infect the host cell and deliver the foreign DNA ([\[link\]](#)). More advanced forms of gene therapy try to correct the mutation at the original site in the genome, such as is the case with treatment of severe combined immunodeficiency (SCID).



Production of Vaccines, Antibiotics, and Hormones

Traditional vaccination strategies use weakened or inactive forms of microorganisms to mount the initial immune response. Modern techniques use the genes of microorganisms cloned into vectors to mass produce the desired antigen. Doctors then introduce the antigen into the body to stimulate the primary immune response and trigger immune memory. The medical field has used genes cloned from the influenza virus to combat the constantly changing strains of this virus.

Antibiotics are a biotechnological product.

Microorganisms, such as fungi, naturally produce them to attain an advantage over bacterial populations. Cultivating and manipulating fungal cells produces antibodies.

Scientists used recombinant DNA technology to produce large-scale quantities of human insulin in *E. coli* as early as 1978. Previously, it was only possible to treat diabetes with pig insulin, which caused allergic reactions in humans because of differences in the gene product. In addition, doctors use human growth hormone (HGH) to treat growth disorders in children. Researchers cloned the HGH gene from a cDNA library and inserted it into *E. coli* cells by cloning it into a bacterial vector.

Transgenic Animals

Although several recombinant proteins in medicine are successfully produced in bacteria, some proteins require a eukaryotic animal host for proper processing. For this reason, the desired genes are cloned and expressed in animals, such as sheep, goats, chickens, and mice. We call animals that have been modified to express recombinant DNA transgenic animals. Several human proteins are expressed in transgenic sheep and goat milk, and some are expressed in chicken eggs. Scientists have used mice extensively for expressing and studying recombinant gene and mutation effects.

Corn, a major agricultural crop used to create products for a variety of industries, is often modified through plant biotechnology. (credit: Keith Weller, USDA)

Transgenic Plants

Manipulating the DNA of plants (i.e., creating GMOs) has helped to create desirable traits, such as disease resistance, herbicide and pesticide resistance, better nutritional value, and better shelf-life ([\[link\]](#)). Plants are the most important source of food for the human population. Farmers developed ways to select for plant varieties with desirable traits long before modern-day biotechnology practices were established.



We call plants that have received recombinant DNA from other species transgenic plants. Because they are not natural, government agencies closely monitor transgenic plants and other GMOs to ensure that they are fit for human consumption and do not endanger other plant and animal life. Because foreign genes can spread to other species in the environment, extensive testing is required to ensure

ecological stability. Staples like corn, potatoes, and tomatoes were the first crop plants that scientists genetically engineered.

Transformation of Plants Using *Agrobacterium tumefaciens*

Gene transfer occurs naturally between species in microbial populations. Many viruses that cause human diseases, such as cancer, act by incorporating their DNA into the human genome. In plants, tumors caused by the bacterium *Agrobacterium tumefaciens* occur by DNA transfer from the bacterium to the plant. Although the tumors do not kill the plants, they stunt the plants and they become more susceptible to harsh environmental conditions. *A. tumefaciens* affects many plants such as walnuts, grapes, nut trees, and beets. Artificially introducing DNA into plant cells is more challenging than in animal cells because of the thick plant cell wall.

Researchers used the natural transfer of DNA from *Agrobacterium* to a plant host to introduce DNA fragments of their choice into plant hosts. In nature, the disease-causing *A. tumefaciens* have a set of plasmids, **Ti plasmids** (tumor-inducing plasmids), that contain genes to produce tumors in plants. DNA from the Ti plasmid integrates into the infected plant cell's genome. Researchers manipulate the Ti plasmids to remove the tumor-causing genes and

insert the desired DNA fragment for transfer into the plant genome. The Ti plasmids carry antibiotic resistance genes to aid selection and researchers can propagate them in *E. coli* cells as well.

The Organic Insecticide *Bacillus thuringiensis*

Bacillus thuringiensis (Bt) is a bacterium that produces protein crystals during sporulation that are toxic to many insect species that affect plants. Insects need to ingest Bt toxin in order to activate the toxin. Insects that have eaten Bt toxin stop feeding on the plants within a few hours. After the toxin activates in the insects' intestines, they die within a couple of days. Modern biotechnology has allowed plants to encode their own crystal Bt toxin that acts against insects. Scientists have cloned the crystal toxin genes from Bt and introduced them into plants. Bt toxin is safe for the environment, nontoxic to humans and other mammals, and organic farmers have approved it as a natural insecticide.

Flavr Savr Tomato

The first GM crop on the market was the Flavr Savr Tomato in 1994. Scientists used antisense RNA technology to slow the softening and rotting process caused by fungal infections, which led to increased shelf life of the GM tomatoes. Additional genetic modification improved the tomato's flavor. The

Flavr Savr tomato did not successfully stay in the market because of problems maintaining and shipping the crop.

Section Summary

Nucleic acids can be isolated from cells for the purposes of further analysis by breaking open the cells and enzymatically destroying all other major macromolecules. Fragmented or whole chromosomes can separate on the basis of size by gel electrophoresis. PCR can amplify short DNA or RNA stretches. Researchers can use Southern and Northern blotting to detect the presence of specific short sequences in a DNA or RNA sample. The term “cloning” may refer to cloning small DNA fragments (molecular cloning), cloning cell populations (cellular cloning), or cloning entire organisms (reproductive cloning). Medical professionals perform genetic testing to identify disease-causing genes, and use gene therapy to cure an inheritable disease.

Transgenic organisms possess DNA from a different species, usually generated by molecular cloning techniques. Vaccines, antibiotics, and hormones are examples of products obtained by recombinant DNA technology. Scientists usually create transgenic plants to improve crop plant characteristics.

Visual Connection Questions

[\[link\]](#) You are working in a molecular biology lab and, unbeknownst to you, your lab partner left the foreign genomic DNA that you are planning to clone on the lab bench overnight instead of storing it in the freezer. As a result, it was degraded by nucleases, but still used in the experiment. The plasmid, on the other hand, is fine. What results would you expect from your molecular cloning experiment?

1. There will be no colonies on the bacterial plate.
2. There will be blue colonies only.
3. There will be blue and white colonies.
4. There will be white colonies only.

[\[link\]](#) B. The experiment would result in blue colonies only.

[\[link\]](#) Do you think Dolly was a Finn-Dorset or a Scottish Blackface sheep?

[\[link\]](#) Dolly was a Finn-Dorset sheep because even though the original cell came from a Scottish blackface sheep and the surrogate

mother was a Scottish blackface, the DNA came from a Finn-Dorset.

Review Questions

GMOs are created by _____.

1. generating genomic DNA fragments with restriction endonucleases
2. introducing recombinant DNA into an organism by any means
3. overexpressing proteins in *E. coli*
4. all of the above

B

Gene therapy can be used to introduce foreign DNA into cells _____.

1. for molecular cloning
2. by PCR
3. of tissues to cure inheritable disease
4. all of the above

C

Insulin produced by molecular cloning:

1. is of pig origin
 2. is a recombinant protein
 3. is made by the human pancreas
 4. is recombinant DNA
-

B

Bt toxin is considered to be _____.

1. a gene for modifying insect DNA
 2. an organic insecticide produced by bacteria
 3. useful for humans to fight against insects
 4. a recombinant protein
-

B

The Flavr Savr Tomato:

1. is a variety of vine-ripened tomato in the supermarket
 2. was created to have better flavor and shelf-life
 3. does not undergo soft rot
 4. all of the above
-

Critical Thinking Questions

Describe the process of Southern blotting.

Southern blotting is the transfer of DNA that has been enzymatically cut into fragments and run on an agarose gel onto a nylon membrane. The DNA fragments that are on the nylon membrane can be denatured to make them single-stranded, and then probed with small DNA fragments that are radioactively or fluorescently labeled, to detect the presence of specific sequences. An example of the use of Southern blotting would be in analyzing the presence, absence, or variation of a disease gene in genomic DNA from a group of patients.

A researcher wants to study cancer cells from a patient with breast cancer. Is cloning the cancer cells an option?

Cellular cloning of the breast cancer cells will establish a cell line, which can be used for

further analysis

How would a scientist introduce a gene for herbicide resistance into a plant?

By identifying an herbicide resistance gene and cloning it into a plant expression vector system, like the Ti plasmid system from *Agrobacterium tumefaciens*. The scientist would then introduce it into the plant cells by transformation, and select cells that have taken up and integrated the herbicide-resistance gene into the genome.

If you had a chance to get your genome sequenced, what are some questions you might be able to have answered about yourself?

What diseases am I prone to and what precautions should I take? Am I a carrier for any disease-causing genes that may be passed on to children?

Glossary

antibiotic resistance

ability of an organism to be unaffected by an antibiotic's actions

biotechnology

use of biological agents for technological advancement

cellular cloning

production of identical cell populations by binary fission

clone

exact replica

foreign DNA

DNA that belongs to a different species or
DNA that is artificially synthesized

gel electrophoresis

technique used to separate molecules on the basis of size using electric charge

gene targeting

method for altering the sequence of a specific gene by introducing the modified version on a vector

gene therapy

technique used to cure inheritable diseases by replacing mutant genes with good genes

genetic diagnosis

diagnosis of the potential for disease development by analyzing disease-causing genes

genetic engineering

alteration of the genetic makeup of an organism

genetic testing

process of testing for the presence of disease-causing genes

genetically modified organism (GMO)

organism whose genome has been artificially changed

host DNA

DNA that is present in the genome of the organism of interest

lysis buffer

solution to break the cell membrane and release cell contents

molecular cloning

cloning of DNA fragments

multiple cloning site (MCS)

site that multiple restriction endonucleases can recognize

Northern blotting

transfer of RNA from a gel to a nylon membrane

polymerase chain reaction (PCR)

technique to amplify DNA

probe

small DNA fragment to determine if the complementary sequence is present in a DNA sample

protease

enzyme that breaks down proteins

recombinant DNA

combining DNA fragments that molecular cloning generates that do not exist in nature; also a chimeric molecule

recombinant protein

a gene's protein product derived by molecular cloning

reproductive cloning

entire organism cloning

restriction endonuclease

enzyme that can recognize and cleave specific DNA sequences

reverse genetics

method of determining the gene's function by starting with the gene itself instead of starting with the gene product

reverse transcriptase PCR (RT-PCR)

PCR technique that involves converting RNA to DNA by reverse transcriptase

ribonuclease

enzyme that breaks down RNA

Southern blotting

DNA transfer from a gel to a nylon membrane

Ti plasmid

plasmid system derived from *Agrobacterium tumifaciens* that scientists have used to introduce foreign DNA into plant cells

transgenic

organism that receives DNA from a different species

Mapping Genomes

By the end of this section, you will be able to do the following:

- Define genomics
- Describe genetic and physical maps
- Describe genomic mapping methods

Genomics is the study of entire genomes, including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species. **Genome mapping** is the process of finding the locations of genes on each chromosome. The maps that genome mapping create are comparable to the maps that we use to navigate streets. A **genetic map** is an illustration that lists genes and their location on a chromosome. Genetic maps provide the big picture (similar to an interstate highway map) and use genetic markers (similar to landmarks). A **genetic marker** is a gene or sequence on a chromosome that co-segregates (shows genetic linkage) with a specific trait. Early geneticists called this linkage analysis. Physical maps present the intimate details of smaller chromosome regions (similar to a detailed road map). A **physical map** is a representation of the physical distance, in nucleotides, between genes or genetic markers. Both genetic linkage maps and physical maps are required to build a genome's complete picture. Having a complete genome map of the genome makes it easier for researchers to

study individual genes. Human genome maps help researchers in their efforts to identify human disease-causing genes related to illnesses like cancer, heart disease, and cystic fibrosis. We can use genome mapping in a variety of other applications, such as using live microbes to clean up pollutants or even prevent pollution. Research involving plant genome mapping may lead to producing higher crop yields or developing plants that better adapt to climate change.

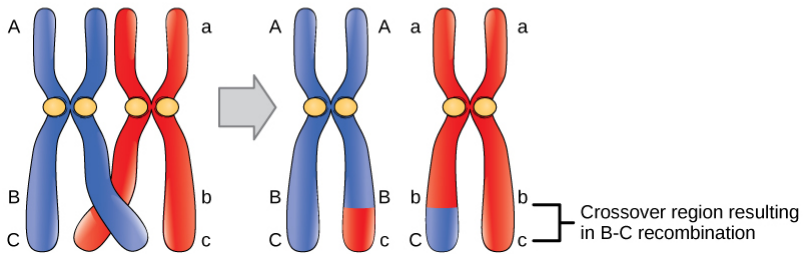
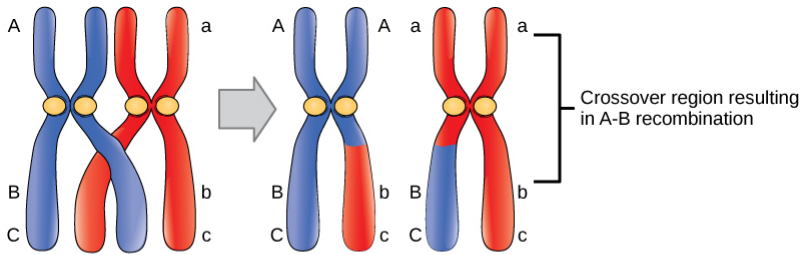
Crossover may occur at different locations on the chromosome. Recombination between genes *A* and *B* is more frequent than recombination between genes *B* and *C* because genes *A* and *B* are farther apart. Therefore, a crossover is more likely to occur between them.

Genetic Maps

The study of genetic maps begins with **linkage analysis**, a procedure that analyzes the recombination frequency between genes to determine if they are linked or show independent assortment. Scientists used the term *linkage* before the discovery of DNA. Early geneticists relied on observing phenotypic changes to understand an organism's genotype. Shortly after Gregor Mendel (the father of modern genetics) proposed that traits were determined by what we now call genes, other researchers observed that different traits were often inherited together, and thereby deduced that the

genes were physically linked by their location on the same chromosome. Gene mapping relative to each other based on linkage analysis led to developing the first genetic maps.

Observations that certain traits were always linked and certain others were not linked came from studying the offspring of crosses between parents with different traits. For example, in garden pea experiments, researchers discovered, that the flower's color and plant pollen's shape were linked traits, and therefore the genes encoding these traits were in close proximity on the same chromosome. We call exchanging DNA between homologous chromosome pairs **genetic recombination**, which occurs by crossing over DNA between homologous DNA strands, such as nonsister chromatids. Linkage analysis involves studying the recombination frequency between any two genes. The greater the distance between two genes, the higher the chance that a recombination event will occur between them, and the higher the recombination frequency between them. [\[link\]](#) shows two possibilities for recombination between two nonsister chromatids during meiosis. If the recombination frequency between two genes is less than 50 percent, they are linked.



The generation of genetic maps requires markers, just as a road map requires landmarks (such as rivers and mountains). Scientists based early genetic maps on using known genes as markers. Scientists now use more sophisticated markers, including those based on non-coding DNA, to compare individuals' genomes in a population. Although individuals of a given species are genetically similar, they are not identical. Every individual has a unique set of traits. These minor differences in the genome between individuals in a population are useful for genetic mapping purposes. In general, a good genetic marker is a region on the chromosome that shows variability or polymorphism (multiple forms) in the population.

Some genetic markers that scientists use in generating genetic maps are **restriction fragment**

length polymorphisms (RFLP), variable number of tandem repeats (VNTRs), **microsatellite polymorphisms**, and the **single nucleotide polymorphisms** (SNPs). We can detect RFLPs (sometimes pronounced “rif-lips”) when the DNA of an individual is cut with a restriction endonuclease that recognizes specific sequences in the DNA to generate a series of DNA fragments, which we can then analyze using gel electrophoresis. Every individual’s DNA will give rise to a unique pattern of bands when cut with a particular set of restriction endonucleases. Scientists sometimes refer to this as an individual’s DNA “fingerprint.” Certain chromosome regions that are subject to polymorphism will lead to generating the unique banding pattern. VNTRs are repeated sets of nucleotides present in DNA’s non-coding regions. Non-coding, or “junk,” DNA has no known biological function; however, research shows that much of this DNA is actually transcribed. While its function is uncertain, it is certainly active, and it may be involved in regulating coding genes. The number of repeats may vary in a population’s individual organisms. Microsatellite polymorphisms are similar to VNTRs, but the repeat unit is very small. SNPs are variations in a single nucleotide.

Because genetic maps rely completely on the natural process of recombination, natural increases or decreases in the recombination level given genome area affects mapping. Some parts of the genome are

recombination hotspots; whereas, others do not show a propensity for recombination. For this reason, it is important to look at mapping information developed by multiple methods.

A cytogenetic map shows the appearance of a chromosome after scientists stain and exam it under a microscope. (credit: National Human Genome Research Institute)

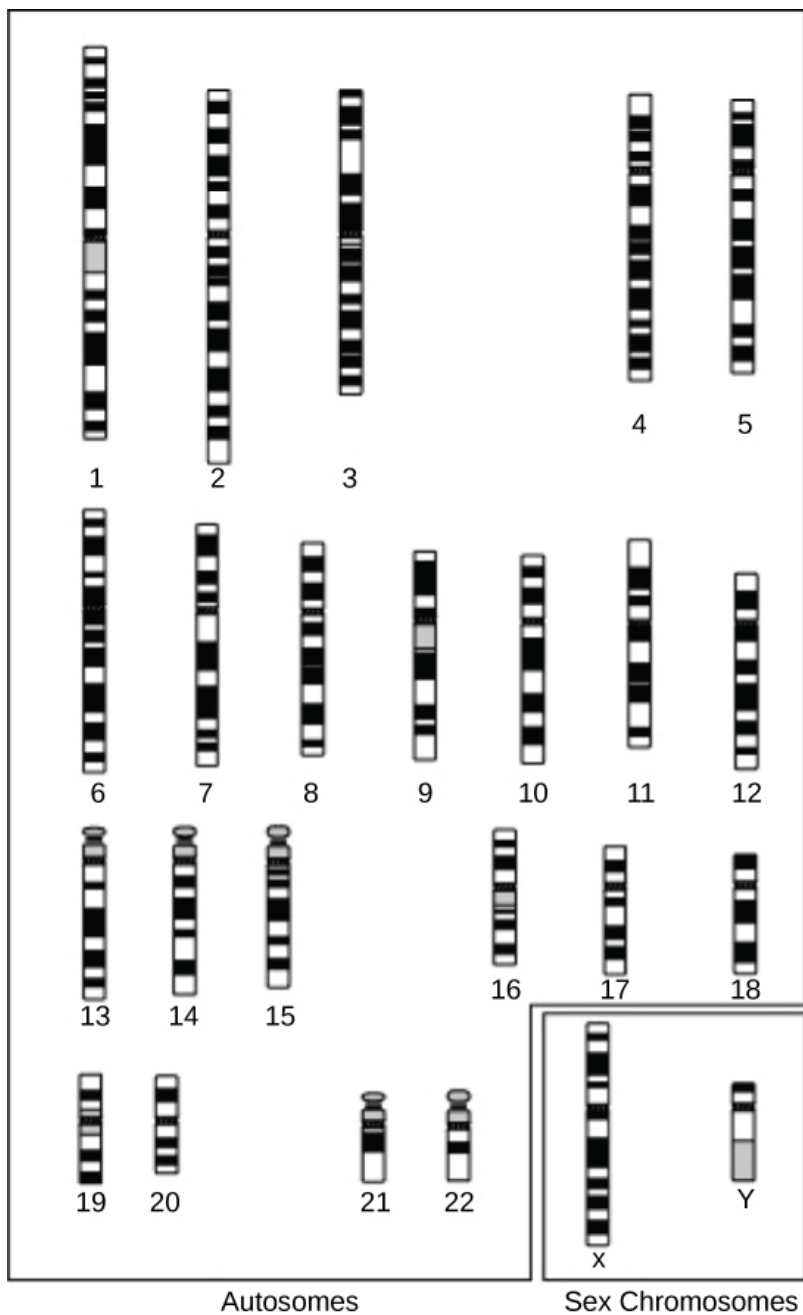
Physical Maps

A physical map provides detail of the actual physical distance between genetic markers, as well as the number of nucleotides. There are three methods scientists use to create a physical map: cytogenetic mapping, radiation hybrid mapping, and sequence mapping. **Cytogenetic mapping** uses information from microscopic analysis of stained chromosome sections ([\[link\]](#)). It is possible to determine the approximate distance between genetic markers using cytogenetic mapping, but not the exact distance (number of base pairs).

Radiation hybrid mapping uses radiation, such as x-rays, to break the DNA into fragments. We can adjust the radiation amount to create smaller or larger fragments. This technique overcomes the limitation of genetic mapping, and we can adjust the radiation so that increased or decreased recombination frequency does not affect it.

Sequence mapping resulted from DNA sequencing technology that allowed for creating detailed

physical maps with distances measured in terms of the number of base pairs. Creating **genomic libraries** and **complementary DNA (cDNA) libraries** (collections of cloned sequences or all DNA from a genome) has sped the physical mapping process. A genetic site that scientists use to generate a physical map with sequencing technology (a sequence-tagged site, or STS) is a unique sequence in the genome with a known exact chromosomal location. An **expressed sequence tag (EST)** and a single sequence length polymorphism (SSLP) are common STSs. An EST is a short STS that we can identify with cDNA libraries, while we obtain SSLPs from known genetic markers, which provide a link between genetic and physical maps.



Genetic and Physical Maps Integration

Genetic maps provide the outline and physical maps provide the details. It is easy to understand why both genome mapping technique types are important to show the big picture. Scientists use information from each technique in combination to study the genome. Scientists are using genomic mapping with different model organisms for research. Genome mapping is still an ongoing process, and as researchers develop more advanced techniques, they expect more breakthroughs. Genome mapping is similar to completing a complicated puzzle using every piece of available data. Mapping information generated in laboratories all over the world goes into central databases, such as GenBank at the National Center for Biotechnology Information (NCBI). Researchers are making efforts for the information to be more easily accessible to other researchers and the general public. Just as we use global positioning systems instead of paper maps to navigate through roadways, NCBI has created a genome viewer tool to simplify the data-mining process.

Scientific Method Connection

How to Use a Genome Map Viewer

Problem statement: Do the human, macaque, and mouse genomes contain common DNA sequences?

Develop a hypothesis.

Go to [this website](#) to test the hypothesis.

The web page displays the comparison of the gene sequences of many organisms to the Human Insulin Receptor gene. Explore the type of information provided, select the groups of organisms needed for testing of the hypothesis from the top portion of the displayed data. Focus the attention to the bottom part, the Selected Orthologues. Explore which columns are relevant to the needed information.

On the same page, there are other options to explore, not all are necessary for the task, however it might give more insight to the value of genome/ gene comparisons.

Link to Learning

Online Mendelian Inheritance in Man (OMIM) is a searchable online catalog of human genes and genetic disorders. This website shows genome mapping information, and also details the history and research of each trait and disorder. Click this [link](#) to search for traits (such as handedness) and genetic disorders (such as diabetes).

Section Summary

Genome mapping is similar to solving a big, complicated puzzle with pieces of information coming from laboratories all over the world. Genetic maps provide an outline for locating genes within a genome, and they estimate the distance between genes and genetic markers on the basis of recombination frequencies during meiosis. Physical maps provide detailed information about the physical distance between the genes. The most detailed information is available through sequence mapping. Researchers combine information from all mapping and sequencing sources to study an entire genome.

Review Questions

ESTs are _____.

1. generated after a cDNA library is made
2. unique sequences in the genome
3. useful for mapping using sequence information
4. all of the above

Linkage analysis _____.

1. is used to create a physical map
 2. is based on the natural recombination process
 3. requires radiation hybrid mapping
 4. involves breaking and rejoining of DNA artificially
-

B

Genetic recombination occurs by which process?

1. independent assortment
 2. crossing over
 3. chromosome segregation
 4. sister chromatids
-

B

Individual genetic maps in a given species are:

1. genetically similar
 2. genetically identical
 3. genetically dissimilar
 4. not useful in species analysis
-

A

Information obtained by microscopic analysis of stained chromosomes is used in:

1. radiation hybrid mapping
2. sequence mapping
3. RFLP mapping
4. cytogenetic mapping

D

Critical Thinking Questions

Why is so much effort being poured into genome mapping applications?

Genome mapping has many different applications and provides comprehensive information that can be used for predictive purposes.

How could a genetic map of the human genome help find a cure for cancer?

A human genetic map can help identify genetic markers and sequences associated with high cancer risk, which can help to screen and provide early detection of different types of cancer.

Glossary

cytogenetic mapping

technique that uses a microscope to create a map from stained chromosomes

expressed sequence tag (EST)

short STS that is identified with cDNA

genetic map

outline of genes and their location on a chromosome

genetic marker

gene or sequence on a chromosome with a known location that is associated with a specific trait

genetic recombination

DNA exchange between homologous chromosome pairs

genome mapping

process of finding the location of genes on each chromosome

cDNA library

collection of cloned cDNA sequences

genomic library

collection of cloned DNA which represents all of the sequences and fragments from a genome

genomics

study of entire genomes including the complete set of genes, their nucleotide sequence and organization, and their interactions within a species and with other species

linkage analysis

procedure that analyzes recombining genes to determine if they are linked

microsatellite polymorphism

variation between individuals in the sequence and number of microsatellite DNA repeats

physical map

representation of the physical distance between genes or genetic markers

radiation hybrid mapping

information obtained by fragmenting the chromosome with x-rays

restriction fragment length polymorphism (RFLP)

variation between individuals in the length of
DNA fragments, which restriction
endonucleases generate

sequence mapping

mapping information obtained after DNA
sequencing

single nucleotide polymorphism (SNP)

variation between individuals in a single
nucleotide

variable number of tandem repeats (VNTRs)

variation in the number of tandem repeats
between individuals in the population

Whole-Genome Sequencing

By the end of this section, you will be able to do the following:

- Describe three types of sequencing
- Define whole-genome sequencing

Although there have been significant advances in the medical sciences in recent years, doctors are still confounded by some diseases, and they are using whole-genome sequencing to discover the root of the problem. **Whole-genome sequencing** is a process that determines an entire genome's DNA sequence. Whole-genome sequencing is a brute-force approach to problem solving when there is a genetic basis at the core of a disease. Several laboratories now provide services to sequence, analyze, and interpret entire genomes.

For example, whole-exome sequencing is a lower-cost alternative to whole genome sequencing. In exome sequencing, the doctor sequences only the DNA's coding, exon-producing regions. In 2010, doctors used whole-exome sequencing to save a young boy whose intestines had multiple mysterious abscesses. The child had several colon operations with no relief. Finally, they performed whole-exome sequencing, which revealed a defect in a pathway that controls apoptosis (programmed cell death). The doctors used a bone-marrow transplant to overcome this genetic disorder, leading to a cure for

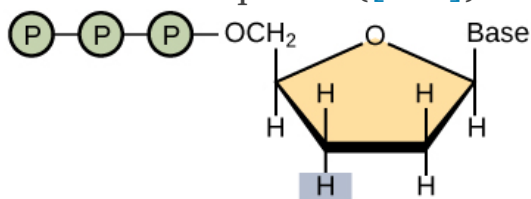
the boy. He was the first person to receive successful treatment based on a whole-exome sequencing diagnosis. Today, human genome sequencing is more readily available and results are available within two days for about \$1000.

A dideoxynucleotide is similar in structure to a deoxynucleotide, but is missing the 3' hydroxyl group (indicated by the box). When a dideoxynucleotide is incorporated into a DNA strand, DNA synthesis stops. This figure illustrates Frederick Sanger's dideoxy chain termination method. Using dideoxynucleotides, the DNA fragment can terminate at different points. The DNA separates on the basis of size, and we can read these bands based on the fragments' size.

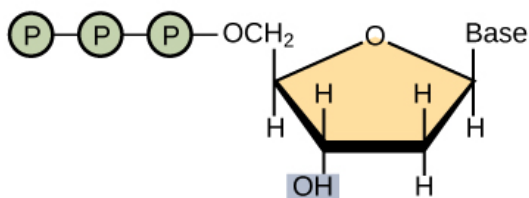
Strategies Used in Sequencing Projects

The basic sequencing technique used in all modern day sequencing projects is the chain termination method (also known as the dideoxy method), which Fred Sanger developed in the 1970s. The chain termination method involves DNA replication of a single-stranded template by using a primer and a regular **deoxynucleotide** (dNTP), which is a monomer, or a single DNA unit. The primer and dNTP mix with a small proportion of fluorescently labeled **dideoxynucleotides** (ddNTPs). The ddNTPs are monomers that are missing a hydroxyl group (–OH) at the site at which another nucleotide usually attaches to form a chain ([\[link\]](#)). Scientists label

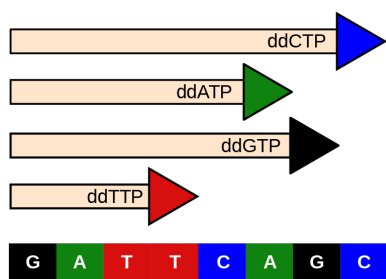
each ddNTP with a different color of fluorophore. Every time a ddNTP incorporates in the growing complementary strand, it terminates the DNA replication process, which results in multiple short strands of replicated DNA that each terminate at a different point during replication. When gel electrophoresis processes the reaction mixture after separating into single strands, the multiple newly replicated DNA strands form a ladder because of the differing sizes. Because the ddNTPs are fluorescently labeled, each band on the gel reflects the DNA strand's size and the ddNTP that terminated the reaction. The different colors of the fluorophore-labeled ddNTPs help identify the ddNTP incorporated at that position. Reading the gel on the basis of each band's color on the ladder produces the template strand's sequence ([\[link\]](#)).



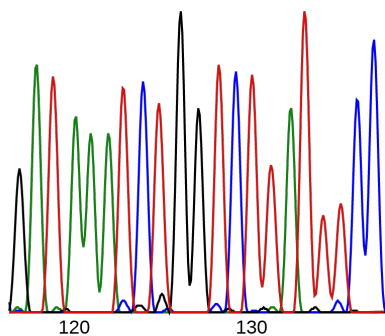
Dideoxynucleotide (ddNTP)



Deoxynucleotide (dNTP)



Dye-labeled dideoxynucleotides are used to generate DNA fragments of different lengths.



G A T A A A T C T G G T C T T A T T T C C

Early Strategies: Shotgun Sequencing and Pair-Wise End Sequencing

In **shotgun sequencing** method, several DNA fragment copies cut randomly into many smaller pieces (somewhat like what happens to a round shot cartridge when fired from a shotgun). All of the segments sequence using the chain-sequencing method. Then, with sequence computer assistance, scientists can analyze the fragments to see where their sequences overlap. By matching overlapping sequences at each fragment's end, scientists can reform the entire DNA sequence. A larger sequence that is assembled from overlapping shorter sequences is called a **contig**. As an analogy, consider that someone has four copies of a landscape photograph that you have never seen before and know nothing about how it should appear. The person then rips up each photograph with their hands, so that different size pieces are present from each copy. The person then mixes all of the pieces together and asks you to reconstruct

the photograph. In one of the smaller pieces you see a mountain. In a larger piece, you see that the same mountain is behind a lake. A third fragment shows only the lake, but it reveals that there is a cabin on the shore of the lake. Therefore, from looking at the overlapping information in these three fragments, you know that the picture contains a mountain behind a lake that has a cabin on its shore. This is the principle behind reconstructing entire DNA sequences using shotgun sequencing.

Originally, shotgun sequencing only analyzed one end of each fragment for overlaps. This was sufficient for sequencing small genomes. However, the desire to sequence larger genomes, such as that of a human, led to developing double-barrel shotgun sequencing, or **pairwise-end sequencing**. In pairwise-end sequencing, scientists analyze each fragment's end for overlap. Pairwise-end sequencing is, therefore, more cumbersome than shotgun sequencing, but it is easier to reconstruct the sequence because there is more available information.

Next-generation Sequencing

Since 2005, automated sequencing techniques used by laboratories are under the umbrella of **next-generation sequencing**, which is a group of automated techniques used for rapid DNA sequencing. These automated low-cost sequencers

can generate sequences of hundreds of thousands or millions of short fragments (25 to 500 base pairs) in the span of one day. These sequencers use sophisticated software to get through the cumbersome process of putting all the fragments in order.

Evolution Connection

Comparing Sequences

A sequence alignment is an arrangement of proteins, DNA, or RNA. Scientists use it to identify similar regions between cell types or species, which may indicate function or structure conservation.

We can use sequence alignments to construct phylogenetic trees. The following website uses a software program called [BLAST \(basic local alignment search tool\)](#).

Under “Basic Blast,” click “Nucleotide Blast.” Input the following sequence into the large “query sequence” box: ATTGCTTCGATTGCA. Below the box, locate the “Species” field and type “human” or “Homo sapiens”. Then click “BLAST” to compare the inputted sequence against the human genome’s known sequences. The result is that this sequence occurs in over a hundred places in the human genome. Scroll down below the graphic with the horizontal bars and you will see a short description of each of the matching hits. Pick one of the hits near the top of the list and click on “Graphics”. This

will bring you to a page that shows the sequence's location within the entire human genome. You can move the slider that looks like a green flag back and forth to view the sequences immediately around the selected gene. You can then return to your selected sequence by clicking the "ATG" button.

Use of Whole-Genome Sequences of Model Organisms

British biochemist and Nobel Prize winner Fred Sanger used a bacterial virus, the bacteriophage *φx174* (5368 base pairs), to completely sequence the first genome. Other scientists later sequenced several other organelle and viral genomes. American biotechnologist, biochemist, geneticist, and businessman Craig Venter sequenced the bacterium *Haemophilus influenzae* in the 1980s. Approximately 74 different laboratories collaborated on sequencing the genome of the yeast *Saccharomyces cerevisiae*, which began in 1989 and was completed in 1996, because it was 60 times bigger than any other genome sequencing. By 1997, the genome sequences of two important model organisms were available: the bacterium *Escherichia coli* K12 and the yeast *Saccharomyces cerevisiae*. We now know the genomes

of other model organisms, such as the mouse *Mus musculus*, the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis. elegans*, and humans *Homo sapiens*. Researchers perform extensive basic research in model organisms because they can apply the information to genetically similar organisms. A **model organism** is a species that researchers use as a model to understand the biological processes in other species that the model organism represents. Having entire genomes sequenced helps with the research efforts in these model organisms. The process of attaching biological information to gene sequences is **genome annotation**. Annotating gene sequences helps with basic experiments in molecular biology, such as designing PCR primers and RNA targets.

Link to Learning

Click through each genome sequencing step at this [site](#).

Genome Sequence Uses

DNA microarrays are methods that scientists use to detect gene expression by analyzing different DNA

fragments that are fixed to a glass slide or a silicon chip to identify active genes and sequences. We can discover almost one million genotypic abnormalities using microarrays; whereas, whole-genome sequencing can provide information about all six billion base pairs in the human genome. Although studying genome sequencing medical applications is interesting, this discipline dwells on abnormal gene function. Knowing about the entire genome will allow researchers to discover future onset diseases and other genetic disorders early. This will allow for more informed decisions about lifestyle, medication, and having children. Genomics is still in its infancy, although someday it may become routine to use whole-genome sequencing to screen every newborn to detect genetic abnormalities.

In addition to disease and medicine, genomics can contribute to developing novel enzymes that convert biomass to biofuel, which results in higher crop and fuel production, and lower consumer cost. This knowledge should allow better methods of control over the microbes that industry uses to produce biofuels. Genomics could also improve monitoring methods that measure the impact of pollutants on ecosystems and help clean up environmental contaminants. Genomics has aided in developing agrochemicals and pharmaceuticals that could benefit medical science and agriculture.

It sounds great to have all the knowledge we can get

from whole-genome sequencing; however, humans have a responsibility to use this knowledge wisely. Otherwise, it could be easy to misuse the power of such knowledge, leading to discrimination based on a person's genetics, human genetic engineering, and other ethical concerns. This information could also lead to legal issues regarding health and privacy.

Section Summary

Whole-genome sequencing is the latest available resource to treat genetic diseases. Some doctors are using whole-genome sequencing to save lives. Genomics has many industrial applications including biofuel development, agriculture, pharmaceuticals, and pollution control. The basic principle of all modern-day sequencing strategies involves the chain termination method of sequencing.

Although the human genome sequences provide key insights to medical professionals, researchers use whole-genome sequences of model organisms to better understand the species' genome. Automation and the decreased cost of whole-genome sequencing may lead to personalized medicine in the future.

Review Questions

The chain termination method of sequencing:

1. uses labeled ddNTPs
 2. uses only dideoxynucleotides
 3. uses only deoxynucleotides
 4. uses labeled dNTPs
-

A

Whole-genome sequencing can be used for advances in:

1. the medical field
 2. agriculture
 3. biofuels
 4. all of the above
-

D

Sequencing an individual person's genome

1. is currently possible
2. could lead to legal issues regarding discrimination and privacy
3. could help make informed choices about medical treatment
4. all of the above

D

What is the most challenging issue facing genome sequencing?

1. the inability to develop fast and accurate sequencing techniques
2. the ethics of using information from genomes at the individual level
3. the availability and stability of DNA
4. all of the above

B

Glossary

chain termination method

method of DNA sequencing using labeled dideoxynucleotides to terminate DNA replication; it is also called the dideoxy method or the Sanger method

contig

larger sequence of DNA assembled from overlapping shorter sequences

deoxynucleotide

individual DNA monomer (single unit)

dideoxynucleotide

individual DNA monomer that is missing a hydroxyl group ($-OH$)

DNA microarray

method to detect gene expression by analyzing many DNA fragments that are fixed to a glass slide or a silicon chip to identify active genes and identify sequences

genome annotation

process of attaching biological information to gene sequences

model organism

species that researchers study and use as a model to understand the biological processes in other species represented by the model organism

next-generation sequencing

group of automated techniques for rapid DNA sequencing

shotgun sequencing

method used to sequence multiple DNA fragments to generate the sequence of a large piece of DNA

whole-genome sequencing

process that determines an entire genome's DNA sequence

Applying Genomics

By the end of this section, you will be able to do the following:

- Explain pharmacogenomics
- Define polygenic

Introducing DNA sequencing and whole genome sequencing projects, particularly the Human Genome project, has expanded the applicability of DNA sequence information. Many fields, such as metagenomics, pharmacogenomics, and mitochondrial genomics are using genomics. Understanding and finding cures for diseases is the most common application of genomics.

Predicting Disease Risk at the Individual Level

Predicting disease risk involves screening currently healthy individuals by genome analysis at the individual level. Health care professionals can recommend intervention with lifestyle changes and drugs before disease onset. However, this approach is most applicable when the problem resides within a single gene defect. Such defects only account for approximately 5 percent of diseases in developed countries. Most of the common diseases, such as heart disease, are multi-factored or **polygenic**,

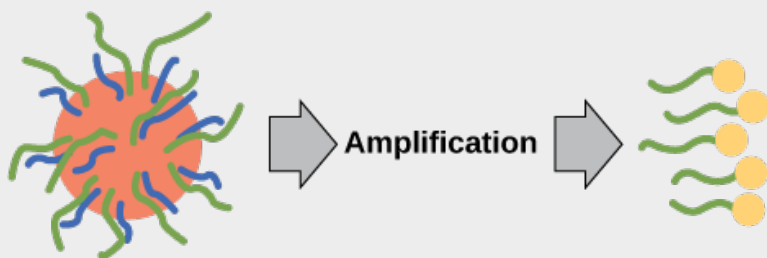
which is a phenotypic characteristic that involves two or more genes, and also involve environmental factors such as diet. In April 2010, scientists at Stanford University published the genome analysis of a healthy individual (Stephen Quake, a scientist at Stanford University, who had his genome sequenced. The analysis predicted his propensity to acquire various diseases. The medical team performed a risk assessment to analyze Quake's percentage of risk for 55 different medical conditions. The team found a rare genetic mutation, which showed him to be at risk for sudden heart attack. The results also predicted that Quake had a 23 percent risk of developing prostate cancer and a 1.4 percent risk of developing Alzheimer's. The scientists used databases and several publications to analyze the genomic data. Even though genomic sequencing is becoming more affordable and analytical tools are becoming more reliable, researchers still must address ethical issues surrounding genomic analysis at a population level.

Visual Connection

PCA3 is a gene that is expressed in prostate epithelial cells and overexpressed in cancerous cells. A high *PCA3* concentration in urine is indicative of prostate cancer. The *PCA3* test is a better indicator of cancer than the more well known PSA test, which measures the level of PSA

(prostate-specific antigen) in the blood.

PCA3



Step 1:

PCA3 mRNA anneals to complementary DNA primers that are attached to beads.

Step 2:

The mRNA is amplified using reverse-transcriptase PCR.

Step 3:

The mRNA is detected using a chemiluminescent probe.

In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The PCA3 test is more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men receive prostate cancer screenings using the PCA3 or PSA test? Should people in general receive screenings to find out if they have a genetic risk for cancer or other diseases?

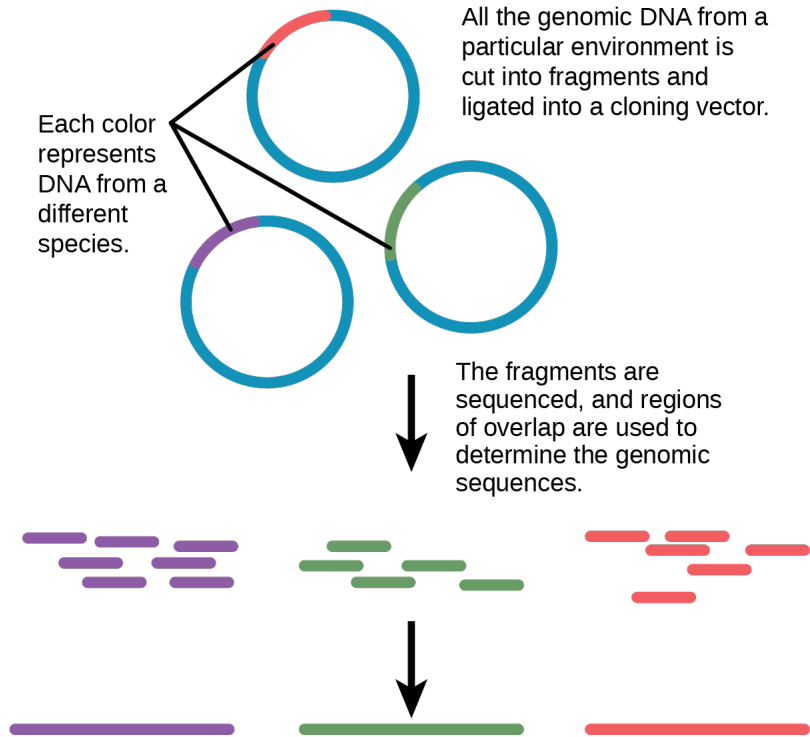
Pharmacogenomics and Toxicogenomics

Pharmacogenomics, or toxicogenomics, involves evaluating drug effectiveness and safety on the basis of information from an individual's genomic sequence. We can study genomic responses to drugs using experimental animals (such as laboratory rats or mice) or live cells in the laboratory before embarking on studies with humans. Studying changes in gene expression could provide information about the transcription profile in the drug's presence, which we can use as an early indicator of the potential for toxic effects. For example, genes involved in cellular growth and controlled cell death, when disturbed, could lead to cancerous cell growth. Genome-wide studies can also help to find new genes involved in drug toxicity. Medical professionals can use personal genome sequence information to prescribe medications that will be most effective and least toxic on the basis of the individual patient's genotype. The gene signatures may not be completely accurate, but medical professionals can test them further before pathologic symptoms arise. Metagenomics involves isolating DNA from multiple species within an environmental niche.

Microbial Genomics: Metagenomics

Traditionally, scholars have taught microbiology with the view that it is best to study microorganisms under **pure culture** conditions. This involves isolating a single cell type and culturing it in the laboratory. Because microorganisms can go through several generations in a matter of hours, their gene expression profiles adapt to the new laboratory environment very quickly. In addition, the vast majority of bacterial species resist culturing in isolation. Most microorganisms do not live as isolated entities, but in microbial communities or biofilms. For all of these reasons, pure culture is not always the best way to study microorganisms.

Metagenomics is the study of the collective genomes of multiple species that grow and interact in an environmental niche. Metagenomics can be used to identify new species more rapidly and to analyze the effect of pollutants on the environment ([link](#)).



Microbial Genomics: Creation of New Biofuels

Knowledge of the genomics of microorganisms is being used to find better ways to harness biofuels from algae and cyanobacteria. The primary sources of fuel today are coal, oil, wood, and other plant products, such as ethanol. Although plants are renewable resources, there is still a need to find more alternative renewable sources of energy to meet our population's energy demands. The microbial world is one of the largest resources for

genes that encode new enzymes and produce new organic compounds, and it remains largely untapped. Microorganisms are used to create products, such as enzymes that are used in research, antibiotics, and other antimicrobial mechanisms. Microbial genomics is helping to develop diagnostic tools, improved vaccines, new disease treatments, and advanced environmental cleanup techniques.

Mitochondrial Genomics

Mitochondria are intracellular organelles that contain their own DNA. Mitochondrial DNA mutates at a rapid rate and scientists often use it to study evolutionary relationships. Another feature that makes studying the mitochondrial genome interesting is that the mitochondrial DNA in most multicellular organisms passes from the mother during the fertilization process. For this reason, scientists often use mitochondrial genomics to trace genealogy.

Experts have used information and clues from DNA samples at crime scenes as evidence in court cases, and they have used genetic markers in forensic analysis. Genomic analysis has also become useful in this field. The first publication showcasing the first use of genomics in forensics came out in 2001. It was a collaborative attempt between academic research institutions and the FBI to solve the

mysterious cases of anthrax communicated via the US Postal Service. Using microbial genomics, researchers determined that the culprit used a specific anthrax strain in all the mailings.

Genomics in Agriculture

Genomics can reduce the trials and failures involved in scientific research to a certain extent, which could improve agricultural crop yield quality and quantity. Linking traits to genes or gene signatures helps improve crop breeding to generate hybrids with the most desirable qualities. Scientists use genomic data to identify desirable traits, and then transfer those traits to a different organism. Researchers are discovering how genomics can improve agricultural production's quality and quantity. For example, scientists could use desirable traits to create a useful product or enhance an existing product, such as making a drought-sensitive crop more tolerant of the dry season.

Section Summary

Imagination is the only barrier to the applicability of genomics. Researchers are applying genomics to most fields of biology. They use it for personalized medicine, prediction of disease risks at an individual

level, studying drug interactions before conducting clinical trials, and studying microorganisms in the environment as opposed to the laboratory. They are also applying it to developments such as generating new biofuels, genealogical assessment using mitochondria, advances in forensic science, and improvements in agriculture.

Visual Connection Questions

[\[link\]](#) In 2011, the United States Preventative Services Task Force recommended against using the PSA test to screen healthy men for prostate cancer. Their recommendation is based on evidence that screening does not reduce the risk of death from prostate cancer. Prostate cancer often develops very slowly and does not cause problems, while the cancer treatment can have severe side effects. The *PCA3* test is considered to be more accurate, but screening may still result in men who would not have been harmed by the cancer itself suffering side effects from treatment. What do you think? Should all healthy men be screened for prostate cancer using the *PCA3* or PSA test? Should people in general be screened to find out if they have a genetic risk for cancer or other diseases?

[\[link\]](#) There are no right or wrong answers to these questions. While it is true that prostate cancer treatment itself can be harmful, many men would rather be aware that they have cancer so they can monitor the disease and begin treatment if it progresses. And while genetic screening may be useful, it is expensive and may cause needless worry. People with certain risk factors may never develop the disease, and preventative treatments may do more harm than good.

Review Questions

Genomics can be used in agriculture to:

1. generate new hybrid strains
2. improve disease resistance
3. improve yield
4. all of the above

D

Genomics can be used on a personal level to:

1. decrease transplant rejection

2. predict genetic diseases that a person may have inherited
 3. determine the risks of genetic diseases for an individual's children
 4. all of the above
-

A

Critical Thinking Questions

Explain why metagenomics is probably the most revolutionary application of genomics.

Metagenomics is revolutionary because it replaced the practice of using pure cultures. Pure cultures were used to study individual species in the laboratory, but did not accurately represent what happens in the environment. Metagenomics studies the genomes of bacterial populations in their environmental niche.

How can genomics be used to predict disease risk and treatment options?

Genomics can provide the unique DNA

sequence of an individual, which can be used for personalized medicine and treatment options.

Glossary

metagenomics

study of multiple species' collective genomes that grow and interact in an environmental niche

pharmacogenomics

study of drug interactions with the genome or proteome; also called toxicogenomics

polygenic

phenotypic characteristic caused by two or more genes

pure culture

growth of a single cell type in the laboratory

Genomics and Proteomics

By the end of this section, you will be able to do the following:

- Explain systems biology
- Describe a proteome
- Define protein signature

Proteins are the final products of genes, which help perform the function that the gene encodes. Amino acids comprise proteins and play important roles in the cell. All enzymes (except ribozymes) are proteins that act as catalysts to affect the rate of reactions. Proteins are also regulatory molecules, and some are hormones. Transport proteins, such as hemoglobin, help transport oxygen to various organs. Antibodies that defend against foreign particles are also proteins. In the diseased state, protein function can be impaired because of changes at the genetic level or because of direct impact on a specific protein.

A **proteome** is the entire set of proteins that a cell type produces. We can study proteomes using the knowledge of genomes because genes code for mRNAs, and the mRNAs encode proteins. Although mRNA analysis is a step in the right direction, not all mRNAs are translated into proteins. **Proteomics** is the study of proteomes' function. Proteomics complements genomics and is useful when scientists want to test their hypotheses that they based on

genes. Even though all multicellular organisms' cells have the same set of genes, the set of proteins produced in different tissues is different and dependent on gene expression. Thus, the genome is constant, but the proteome varies and is dynamic within an organism. In addition, RNAs can be alternately spliced (cut and pasted to create novel combinations and novel proteins) and many proteins modify themselves after translation by processes such as proteolytic cleavage, phosphorylation, glycosylation, and ubiquitination. There are also protein-protein interactions, which complicate studying proteomes. Although the genome provides a blueprint, the final architecture depends on several factors that can change the progression of events that generate the proteome.

Metabolomics is related to genomics and proteomics. **Metabolomics** involves studying small molecule metabolites in an organism. The **metabolome** is the complete set of metabolites that are related to an organism's genetic makeup. Metabolomics offers an opportunity to compare genetic makeup and physical characteristics, as well as genetic makeup and environmental factors. The goal of metabolome research is to identify, quantify, and catalogue all the metabolites in living organisms' tissues and fluids.

Scientists use two-hybrid screening to determine whether two proteins interact. In this method, a transcription factor splits into a DNA-binding

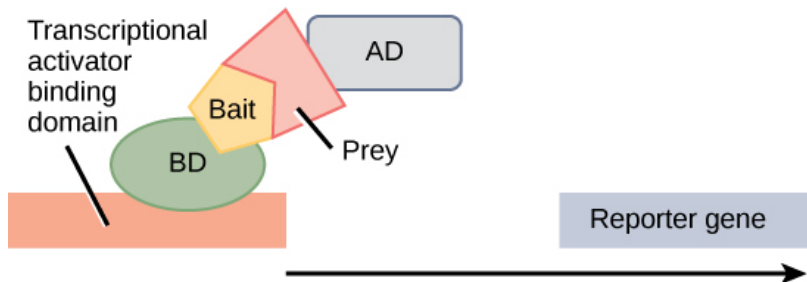
domain (BD) and an activator domain (AD). The binding domain is able to bind the promoter in the activator domain's absence, but it does not turn on transcription. The bait protein attaches to the BD, and the prey protein attaches to the AD. Transcription occurs only if the prey “catches” the bait.

Basic Techniques in Protein Analysis

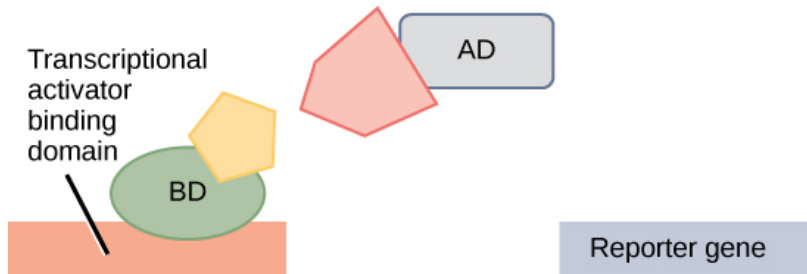
The ultimate goal of proteomics is to identify or compare the proteins expressed from a given genome under specific conditions, study the interactions between the proteins, and use the information to predict cell behavior or develop drug targets. Just as scientists analyze the genome using the basic DNA sequencing technique, proteomics requires techniques for protein analysis. The basic technique for protein analysis, analogous to DNA sequencing, is mass spectrometry. Mass spectrometry identifies and determines a molecule's characteristics. Advances in spectrometry have allowed researchers to analyze very small protein samples. X-ray crystallography, for example, enables scientists to determine a protein crystal's three-dimensional structure at atomic resolution. Another protein imaging technique, nuclear magnetic resonance (NMR), uses atoms' magnetic properties to determine the protein's three-dimensional structure in aqueous solution. Scientists have also used protein microarrays to study protein

interactions. Large-scale adaptations of the basic two-hybrid screen ([\[link\]](#)) have provided the basis for protein microarrays. Scientists use computer software to analyze the vast amount of data for proteomic analysis.

Genomic- and proteomic-scale analyses are part of **systems biology**, which is the study of whole biological systems (genomes and proteomes) based on interactions within the system. The European Bioinformatics Institute and the Human Proteome Organization (HUPO) are developing and establishing effective tools to sort through the enormous pile of systems biology data. Because proteins are the direct products of genes and reflect activity at the genomic level, it is natural to use proteomes to compare the protein profiles of different cells to identify proteins and genes involved in disease processes. Most pharmaceutical drug trials target proteins. Researchers use information that they obtain from proteomics to identify novel drugs and to understand their mechanisms of action.



If the bait protein interacts with the prey protein, the promoter's activator domain binds to the binding domain, and transcription occurs.



If the prey doesn't catch the bait no transcription occurs.

Scientists are challenged when implementing proteomic analysis because it is difficult to detect small protein quantities. Although mass spectrometry is good for detecting small protein amounts, variations in protein expression in diseased states can be difficult to discern. Proteins are naturally unstable molecules, which makes proteomic analysis much more difficult than genomic analysis.

Cancer Proteomics

Researchers are studying patients' genomes and proteomes to understand the genetic basis of diseases. The most prominent disease researchers are studying with proteomic approaches is cancer. These approaches improve screening and early cancer detection. Researchers are able to identify proteins whose expression indicates the disease process. An individual protein is a **biomarker**; whereas, a set of proteins with altered expression levels is a **protein signature**. For a biomarker or protein signature to be useful as a candidate for early cancer screening and detection, they must secrete in body fluids, such as sweat, blood, or urine, such that health professionals can perform large-scale screenings in a noninvasive fashion. The current problem with using biomarkers for early cancer detection is the high rate of false-negative results. A **false negative** is an incorrect test result that should have been positive. In other words, many cancer cases go undetected, which makes biomarkers unreliable. Some examples of protein biomarkers in cancer detection are CA-125 for ovarian cancer and PSA for prostate cancer. Protein signatures may be more reliable than biomarkers to detect cancer cells. Researchers are also using proteomics to develop individualized treatment plans, which involves predicting whether or not an individual will respond to specific drugs and the side effects that the individual may experience. Researchers also use proteomics to predict the possibility of disease recurrence.

The National Cancer Institute has developed programs to improve cancer detection and treatment. The Clinical Proteomic Technologies for Cancer and the Early Detection Research Network are efforts to identify protein signatures specific to different cancer types. The Biomedical Proteomics Program identifies protein signatures and designs effective therapies for cancer patients.

Section Summary

Proteomics is the study of the entire set of proteins expressed by a given type of cell under certain environmental conditions. In a multicellular organism, different cell types will have different proteomes, and these will vary with environmental changes. Unlike a genome, a proteome is dynamic and in constant flux, which makes it both more complicated and more useful than the knowledge of genomes alone.

Proteomics approaches rely on protein analysis. Researchers are constantly upgrading these techniques. Researchers have used proteomics to study different cancer types. Medical professionals are using different biomarkers and protein signatures to analyze each cancer type. The future goal is to have a personalized treatment plan for each individual.

Review Questions

What is a biomarker?

1. the color coding of different genes
 2. a protein that is uniquely produced in a diseased state
 3. a molecule in the genome or proteome
 4. a marker that is genetically inherited
-

B

A protein signature is:

1. the path followed by a protein after it is synthesized in the nucleus
 2. the path followed by a protein in the cytoplasm
 3. a protein expressed on the cell surface
 4. a unique set of proteins present in a diseased state
-

D

Critical Thinking Questions

How has proteomics been used in cancer detection and treatment?

Proteomics has provided a way to detect biomarkers and protein signatures, which have been used to screen for the early detection of cancer.

What is personalized medicine?

Personalized medicine is the use of an individual's genomic sequence to predict the risk for specific diseases. When a disease does occur, it can be used to develop a personalized treatment plan.

Glossary

biomarker

individual protein that is uniquely produced in a diseased state

false negative

incorrect test result that should have been positive

metabolome

complete set of metabolites which are related to an organism's genetic makeup

metabolomics

study of small molecule metabolites in an organism

protein signature

set of uniquely expressed proteins in the diseased state

proteome

entire set of proteins that cell type produces

proteomics

study of proteomes' function

systems biology

study of whole biological systems (genomes and proteomes) based on interactions within the system

Introduction

class = "introduction" All organisms are products of evolution adapted to their environment. (a) Saguaro (*Carnegiea gigantea*) can soak up 750 liters of water in a single rain storm, enabling these cacti to survive the dry conditions of the Sonora desert in Mexico and the Southwestern United States. (b) The Andean semiaquatic lizard (*Potamites montanicola*) discovered in Peru in 2010 lives between 1,570 to 2,100 meters in elevation, and, unlike most lizards, is nocturnal and swims. Scientists still do not know how these cold-blood animals are able to move in the cold (10 to 15°C) temperatures of the Andean night. (credit a: modification of work by Gentry George, U.S. Fish and Wildlife Service; credit b: modification of work by Germán Chávez and Diego Vásquez, ZooKeys)



All living organisms, from bacteria to baboons to blueberries, evolved at some point from a different species. Although it may seem that living things today stay much the same, that is not the case—evolution is an ongoing process.

The theory of evolution is the unifying theory of biology, meaning it is the framework within which biologists ask questions about the living world. Its power is that it provides direction for predictions about living things that are borne out in ongoing experiments. The Ukrainian-born American geneticist Theodosius Dobzhansky famously wrote that “nothing makes sense in biology except in the light of evolution.”[\[footnote\]](#) He meant that the tenet that all life has evolved and diversified from a common ancestor is the foundation from which we approach all questions in biology.

Theodosius Dobzhansky. “Biology, Molecular and Organismic.” *American Zoologist* 4, no. 4 (1964): 449.

Understanding Evolution

By the end of this section, you will be able to do the following:

- Describe how scientists developed the present-day theory of evolution
- Define adaptation
- Explain convergent and divergent evolution
- Describe homologous and vestigial structures
- Discuss misconceptions about the theory of evolution

Evolution by natural selection describes a mechanism for how species change over time. Scientists, philosophers, researchers, and others had made suggestions and debated this topic well before Darwin began to explore this idea. Classical Greek philosopher Plato emphasized in his writings that species were static and unchanging, yet there were also ancient Greeks who expressed evolutionary ideas. In the eighteenth century, naturalist Georges-Louis Leclerc Comte de Buffon reintroduced ideas about the evolution of animals and observed that various geographic regions have different plant and animal populations, even when the environments are similar. Some at this time also accepted that there were extinct species.

Also during the eighteenth century, James Hutton, a Scottish geologist and naturalist, proposed that geological change occurred gradually by

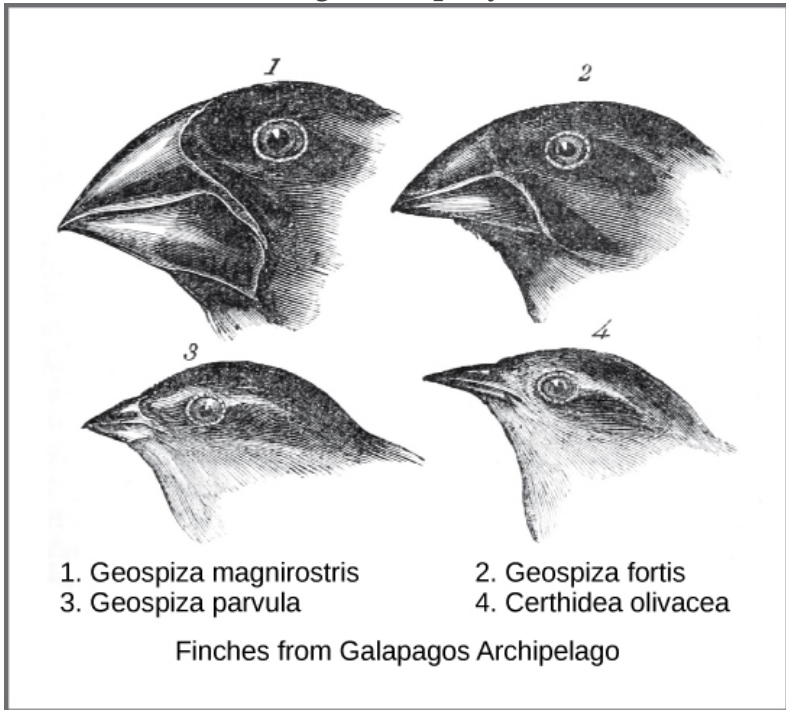
accumulating small changes from processes operating like they are today over long periods of time. This contrasted with the predominant view that the planet's geology was a consequence of catastrophic events occurring during a relatively brief past. Nineteenth century geologist Charles Lyell popularized Hutton's view. A friend to Darwin, Lyell's ideas were influential on Darwin's thinking: Lyell's notion of the greater age of Earth gave more time for gradual change in species, and the process of change provided an analogy for this change. In the early nineteenth century, Jean-Baptiste Lamarck published a book that detailed a mechanism for evolutionary change. We now refer to this mechanism as an inheritance of acquired characteristics by which the environment causes modifications in an individual, or offspring could use or disuse of a structure during its lifetime, and thus bring about change in a species. While many discredited this mechanism for evolutionary change, Lamarck's ideas were an important influence on evolutionary thought.

Darwin observed that beak shape varies among finch species. He postulated that ancestral species' beaks had adapted over time to equip the finches to acquire different food sources. Both (a) Charles Darwin and (b) Alfred Wallace wrote scientific papers on natural selection that they presented together at the Linnean Society in 1858.

Charles Darwin and Natural Selection

In the mid-nineteenth century, two naturalists, Charles Darwin and Alfred Russel Wallace, independently conceived and described the actual mechanism for evolution. Importantly, each naturalist spent time exploring the natural world on expeditions to the tropics. From 1831 to 1836, Darwin traveled around the world on *H.M.S. Beagle*, including stops in South America, Australia, and the southern tip of Africa. Wallace traveled to Brazil to collect insects in the Amazon rainforest from 1848 to 1852 and to the Malay Archipelago from 1854 to 1862. Darwin's journey, like Wallace's later journeys to the Malay Archipelago, included stops at several island chains, the last being the Galápagos Islands west of Ecuador. On these islands, Darwin observed species of organisms on different islands that were clearly similar, yet had distinct differences. For example, the ground finches inhabiting the Galápagos Islands comprised several species with a unique beak shape ([link](#)). The species on the islands had a graded series of beak sizes and shapes with very small differences between the most similar. He observed that these finches closely resembled another finch species on the South American mainland. Darwin imagined that the island species might be species modified from one of the original mainland species. Upon further study, he realized that each finch's varied beaks helped the birds acquire a specific type of food. For example, seed-eating finches had stronger, thicker beaks for breaking seeds, and insect-eating finches had spear-

like beaks for stabbing their prey.



Wallace and Darwin both observed similar patterns in other organisms and they independently developed the same explanation for how and why such changes could take place. Darwin called this mechanism natural selection. **Natural selection**, or “survival of the fittest,” is the more prolific reproduction of individuals with favorable traits that survive environmental change because of those traits. This leads to evolutionary change.

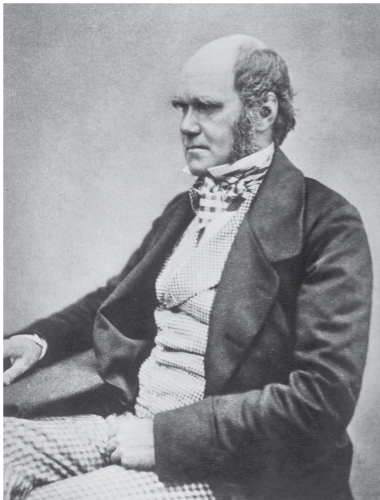
For example, Darwin observed a population of giant tortoises in the Galápagos Archipelago to have longer necks than those that lived on other islands with dry lowlands. These tortoises were “selected”

because they could reach more leaves and access more food than those with short necks. In times of drought when fewer leaves would be available, those that could reach more leaves had a better chance to eat and survive than those that couldn't reach the food source. Consequently, long-necked tortoises would be more likely to be reproductively successful and pass the long-necked trait to their offspring. Over time, only long-necked tortoises would be present in the population.

Natural selection, Darwin argued, was an inevitable outcome of three principles that operated in nature. First, most characteristics of organisms are inherited, or passed from parent to offspring. Although no one, including Darwin and Wallace, knew how this happened at the time, it was a common understanding. Second, more offspring are produced than are able to survive, so resources for survival and reproduction are limited. The capacity for reproduction in all organisms outstrips the availability of resources to support their numbers. Thus, there is competition for those resources in each generation. Both Darwin and Wallace's understanding of this principle came from reading economist Thomas Malthus' essay that explained this principle in relation to human populations. Third, offspring vary among each other in regard to their characteristics and those variations are inherited. Darwin and Wallace reasoned that offspring with inherited characteristics which allow

them to best compete for limited resources will survive and have more offspring than those individuals with variations that are less able to compete. Because characteristics are inherited, these traits will be better represented in the next generation. This will lead to change in populations over generations in a process that Darwin called descent with modification. Ultimately, natural selection leads to greater adaptation of the population to its local environment. It is the only mechanism known for adaptive evolution.

In 1858, Darwin and Wallace ([\[link\]](#)) presented papers at the Linnean Society in London that discussed the idea of natural selection. The following year Darwin's book, *On the Origin of Species*, was published. His book outlined in considerable detail his arguments for evolution by natural selection.



(a)



(b)

It is difficult and time-consuming to document and present examples of evolution by natural selection. The Galápagos finches are an excellent example. Peter and Rosemary Grant and their colleagues have studied Galápagos finch populations every year since 1976 and have provided important evidence of natural selection. The Grants found changes from one generation to the next in beak shape distribution with the medium ground finch on the Galápagos island of Daphne Major. The birds have inherited a variation in their bill shape with some having wide deep bills and others having thinner bills. During a period in which rainfall was higher than normal because of an El Niño, there was a lack of large hard seeds of which the large-billed birds ate; however, there was an abundance of the small soft seeds which the small-billed birds ate. Therefore, the small-billed birds were able to survive and reproduce. In the years following this El Niño, the Grants measured beak sizes in the population and found that the average bill size was smaller. Since bill size is an inherited trait, parents with smaller bills had more offspring and the bill evolved into a much smaller size. As conditions improved in 1987 and larger seeds became more available, the trend toward smaller average bill size ceased.

Career Connection

Field Biologist

Many people hike, explore caves, scuba dive, or climb mountains for recreation. People often participate in these activities hoping to see wildlife. Experiencing the outdoors can be incredibly enjoyable and invigorating. What if your job entailed working in the wilderness? Field biologists by definition work outdoors in the “field.” The term field in this case refers to any location outdoors, even under water. A field biologist typically focuses research on a certain species, group of organisms, or a single habitat ([\[link\]](#)). A field biologist tranquilizes a polar bear for study. (credit: Karen Rhode)



One objective of many field biologists includes discovering new, unrecorded species. Not only do such findings expand our understanding of the natural world, but they also lead to important

innovations in fields such as medicine and agriculture. Plant and microbial species, in particular, can reveal new medicinal and nutritive knowledge. Other organisms can play key roles in ecosystems or if rare require protection. When discovered, researchers can use these important species as evidence for environmental regulations and laws.

Flowering plants evolved from a common ancestor. Notice that the (a) dense blazing star (*Liatrus spicata*) and the (b) purple coneflower (*Echinacea purpurea*) vary in appearance, yet both share a similar basic morphology. (credit a: modification of work by Drew Avery; credit b: modification of work by Cory Zanker)

Processes and Patterns of Evolution

Natural selection can only take place if there is **variation**, or differences, among individuals in a population. Importantly, these differences must have some genetic basis; otherwise, the selection will not lead to change in the next generation. This is critical because nongenetic reasons can cause variation among individuals such as an individual's height because of better nutrition rather than different genes.

Genetic diversity in a population comes from two main mechanisms: mutation and sexual reproduction. Mutation, a change in DNA, is the ultimate source of new alleles, or new genetic variation in any population. The genetic changes that mutation causes can have one of three outcomes on the phenotype. A mutation affects the organism's phenotype in a way that gives it reduced fitness—lower likelihood of survival or fewer offspring. A mutation may produce a phenotype with a beneficial effect on fitness. Many mutations will also have no effect on the phenotype's fitness. We call these neutral mutations. Mutations may also have a whole range of effect sizes on the organism's fitness that expresses them in their phenotype, from a small effect to a great effect. Sexual reproduction also leads to genetic diversity: when two parents reproduce, unique combinations of alleles assemble to produce the unique genotypes and thus phenotypes in each offspring.

We call a heritable trait that helps an organism's survival and reproduction in its present environment an **adaptation**. Scientists describe groups of organisms adapting to their environment when a genetic variation occurs over time that increases or maintains the population's “fit” to its environment. A platypus's webbed feet are an adaptation for swimming. A snow leopard's thick fur is an adaptation for living in the cold. A cheetah's fast speed is an adaptation for catching prey.

Whether or not a trait is favorable depends on the current environmental conditions. The same traits are not always selected because environmental conditions can change. For example, consider a plant species that grew in a moist climate and did not need to conserve water. Large leaves were selected because they allowed the plant to obtain more energy from the sun. Large leaves require more water to maintain than small leaves, and the moist environment provided favorable conditions to support large leaves. After thousands of years, the climate changed, and the area no longer had excess water. The direction of natural selection shifted so that plants with small leaves were selected because those populations were able to conserve water to survive the new environmental conditions.

The evolution of species has resulted in enormous variation in form and function. Sometimes, evolution gives rise to groups of organisms that become tremendously different from each other. We call two species that evolve in diverse directions from a common point **divergent evolution**. We can see such divergent evolution in the forms of the reproductive organs of flowering plants which share the same basic anatomies; however, they can look very different as a result of selection in different physical environments and adaptation to different kinds of pollinators ([link](#)).



(a)



(b)

In other cases, similar phenotypes evolve independently in distantly related species. For example, flight has evolved in both bats and insects, and they both have structures we refer to as wings, which are adaptations to flight. However, bat and insect wings have evolved from very different original structures. We call this phenomenon **convergent evolution**, where similar traits evolve independently in species that do not share a common ancestry. The two species came to the same function, flying, but did so separately from each other.

These physical changes occur over enormous time spans and help explain how evolution occurs. Natural selection acts on individual organisms, which can then shape an entire species. Although natural selection may work in a single generation on an individual, it can take thousands or even millions of years for an entire species' genotype to evolve. It is over these large time spans that life on earth has changed and continues to change.

In this (a) display, fossil hominids are arranged from oldest (bottom) to newest (top). As hominids evolved, the skull's shape changed. An artist's rendition of (b) extinct species of the genus *Equus* reveals that these ancient species resembled the modern horse (*Equus ferus*) but varied in size. The similar construction of these appendages indicates that these organisms share a common ancestor. The white winter coat of the (a) arctic fox and the (b) ptarmigan's plumage are adaptations to their environments. (credit a: modification of work by Keith Morehouse)

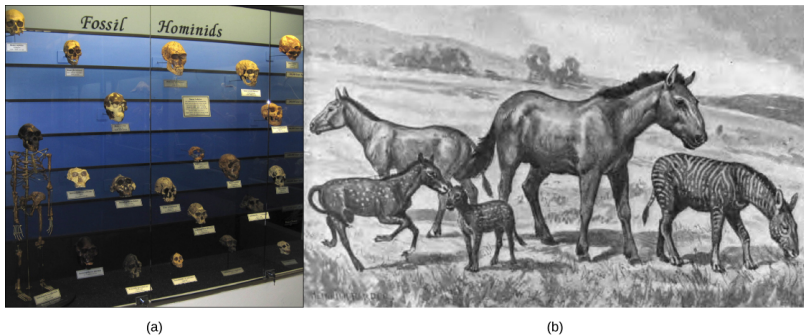
Evidence of Evolution

The evidence for evolution is compelling and extensive. Looking at every level of organization in living systems, biologists see the signature of past and present evolution. Darwin dedicated a large portion of his book, *On the Origin of Species*, to identifying patterns in nature that were consistent with evolution, and since Darwin, our understanding has become clearer and broader.

Fossils

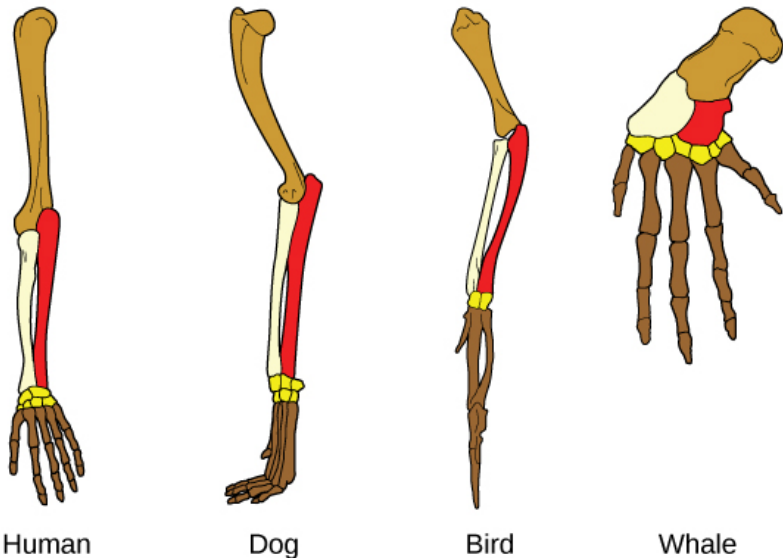
Fossils provide solid evidence that organisms from the past are not the same as those today, and fossils show a progression of evolution. Scientists determine the age of fossils and categorize them

from all over the world to determine when the organisms lived relative to each other. The resulting fossil record tells the story of the past and shows the evolution of form over millions of years ([link]). For example, scientists have recovered highly detailed records showing the evolution of humans and horses ([link]). The whale flipper shares a similar morphology to bird and mammal appendages ([link]) indicating that these species share a common ancestor.



Anatomy and Embryology

Another type of evidence for evolution is the presence of structures in organisms that share the same basic form. For example, the bones in human, dog, bird, and whale appendages all share the same overall construction ([link]) resulting from their origin in a common ancestor's appendages. Over time, evolution led to changes in the bones' shapes and sizes different species, but they have maintained the same overall layout. Scientists call these synonymous parts **homologous structures**.



Some structures exist in organisms that have no apparent function at all, and appear to be residual parts from a past common ancestor. We call these unused structures without function **vestigial structures**. Other examples of vestigial structures are wings on flightless birds, leaves on some cacti, and hind leg bones in whales.

Link to Learning

Visit this [interactive site](#) to guess which bone structures are homologous and which are analogous, and see examples of evolutionary adaptations to illustrate these concepts.

Another evidence of evolution is the convergence of form in organisms that share similar environments. For example, species of unrelated animals, such as the arctic fox and ptarmigan, living in the arctic region have been selected for seasonal white phenotypes during winter to blend with the snow and ice ([link](#)). These similarities occur not because of common ancestry, but because of similar selection pressures—the benefits of predators not seeing them.



(a)

(b)

Embryology, the study of the anatomy of an organism's development to its adult form, also provides evidence of relatedness between now widely divergent groups of organisms. Mutational tweaking in the embryo can have such magnified consequences in the adult that tends to conserve embryo formation. As a result, structures that are absent in some groups often appear in their embryonic forms and disappear when they reach the

adult or juvenile form. For example, all vertebrate embryos, including humans, exhibit gill slits and tails at some point in their early development. These disappear in the adults of terrestrial groups but adult forms of aquatic groups such as fish and some amphibians maintain them. Great ape embryos, including humans, have a tail structure during their development that they lose when they are born.

Biogeography

The geographic distribution of organisms on the planet follows patterns that we can explain best by evolution in conjunction with tectonic plate movement over geological time. Broad groups that evolved before the supercontinent Pangaea broke up (about 200 million years ago) are distributed worldwide. Groups that evolved since the breakup appear uniquely in regions of the planet, such as the unique flora and fauna of northern continents that formed from the supercontinent Laurasia and of the southern continents that formed from the supercontinent Gondwana. The presence of members of the plant family Proteaceae in Australia, southern Africa, and South America was most predominant prior to the southern supercontinent Gondwana breaking up.

Marsupial diversification in Australia and the absence of other mammals reflect Australia's long isolation. Australia has an abundance of endemic

species—species found nowhere else—which is typical of islands whose isolation by expanses of water prevents species to migrate. Over time, these species diverge evolutionarily into new species that look very different from their ancestors that may exist on the mainland. Australia's marsupials, the Galápagos' finches, and many species on the Hawaiian Islands are all unique to their one point of origin, yet they display distant relationships to ancestral species on mainlands.

Molecular Biology

Like anatomical structures, the molecular structures of life reflect descent with modification. DNA's universality reflects evidence of a common ancestor for all of life. Fundamental divisions in life between the genetic code, DNA replication, and expression are reflected in major structural differences in otherwise conservative structures such as ribosome components and membrane structures. In general, the relatedness of groups of organisms is reflected in the similarity of their DNA sequences—exactly the pattern that we would expect from descent and diversification from a common ancestor.

DNA sequences have also shed light on some of the mechanisms of evolution. For example, it is clear that the evolution of new functions for proteins commonly occurs after gene duplication events that allow freely modifying one copy by mutation,

selection, or drift (changes in a population's gene pool resulting from chance), while the second copy continues to produce a functional protein.

Misconceptions of Evolution

Although the theory of evolution generated some controversy when Darwin first proposed it, biologists almost universally accepted it, particularly younger biologists, within 20 years after publication of *On the Origin of Species*. Nevertheless, the theory of evolution is a difficult concept and misconceptions about how it works abound.

Link to Learning

This [site](#) addresses some of the main misconceptions associated with the theory of evolution.

Evolution Is Just a Theory

Critics of the theory of evolution dismiss its importance by purposefully confounding the everyday usage of the word “theory” with the way scientists use the word. In science, we understand a “theory” to be a body of thoroughly tested and

verified explanations for a set of observations of the natural world. Scientists have a theory of the atom, a theory of gravity, and the theory of relativity, each which describes understood facts about the world. In the same way, the theory of evolution describes facts about the living world. As such, a theory in science has survived significant efforts to discredit it by scientists. In contrast, a “theory” in common vernacular is a word meaning a guess or suggested explanation. This meaning is more akin to the scientific concept of “hypothesis.” When critics of evolution say it is “just a theory,” they are implying that there is little evidence supporting it and that it is still in the process of rigorous testing. This is a mischaracterization.

Individuals Evolve

Evolution is the change in a population's genetic composition over time, specifically over generations, resulting from differential reproduction of individuals with certain alleles. Individuals do change over their lifetime, obviously, but this is development and involves changes programmed by the set of genes the individual acquired at birth in coordination with the individual's environment. When thinking about the evolution of a characteristic, it is probably best to think about the change of the average value of the characteristic in the population over time. For example, when natural selection leads to bill-size change in medium

ground finches in the Galápagos, this does not mean that individual bills on the finches are changing. If one measures the average bill size among all individuals in the population at one time and then measures them in the population several years later, this average value will be different as a result of evolution. Although some individuals may survive from the first time to the second, they will still have the same bill size; however, there will be many new individuals who contribute to the shift in average bill size.

Evolution Explains the Origin of Life

It is a common misunderstanding that evolution includes an explanation of life's origins. Conversely, some of the theory's critics believe that it cannot explain the origin of life. The theory does not try to explain the origin of life. The theory of evolution explains how populations change over time and how life diversifies the origin of species. It does not shed light on the beginnings of life including the origins of the first cells, which define life. Importantly, biologists believe that the presence of life on Earth precludes the possibility that the events that led to life on Earth can repeat themselves because the intermediate stages would immediately become food for existing living things.

However, once a mechanism of inheritance was in place in the form of a molecule like DNA either

within a cell or pre-cell, these entities would be subject to the principle of natural selection. More effective reproducers would increase in frequency at the expense of inefficient reproducers. While evolution does not explain the origin of life, it may have something to say about some of the processes operating once pre-living entities acquired certain properties.

Organisms Evolve on Purpose

Statements such as “organisms evolve in response to a change in an environment” are quite common, but such statements can lead to two types of misunderstandings. First, do not interpret the statement to mean that individual organisms evolve. The statement is shorthand for “a population evolves in response to a changing environment.” However, a second misunderstanding may arise by interpreting the statement to mean that the evolution is somehow intentional. A changed environment results in some individuals in the population, those with particular phenotypes, benefiting and therefore producing proportionately more offspring than other phenotypes. This results in change in the population if the characteristics are genetically determined.

It is also important to understand that the variation that natural selection works on is already in a population and does not arise in response to an

environmental change. For example, applying antibiotics to a population of bacteria will, over time, select a population of bacteria that are resistant to antibiotics. The resistance, which a gene causes, did not arise by mutation because of applying the antibiotic. The gene for resistance was already present in the bacteria's gene pool, likely at a low frequency. The antibiotic, which kills the bacterial cells without the resistance gene, strongly selects individuals that are resistant, since these would be the only ones that survived and divided. Experiments have demonstrated that mutations for antibiotic resistance do not arise as a result of antibiotic.

In a larger sense, evolution is not goal directed. Species do not become “better” over time. They simply track their changing environment with adaptations that maximize their reproduction in a particular environment at a particular time. Evolution has no goal of making faster, bigger, more complex, or even smarter species, despite the commonness of this kind of language in popular discourse. What characteristics evolve in a species are a function of the variation present and the environment, both of which are constantly changing in a nondirectional way. A trait that fits in one environment at one time may well be fatal at some point in the future. This holds equally well for insect and human species.

Section Summary

Evolution is the process of adaptation through mutation which allows more desirable characteristics to pass to the next generation. Over time, organisms evolve more characteristics that are beneficial to their survival. For living organisms to adapt and change to environmental pressures, genetic variation must be present. With genetic variation, individuals have differences in form and function that allow some to survive certain conditions better than others. These organisms pass their favorable traits to their offspring. Eventually, environments change, and what was once a desirable, advantageous trait may become an undesirable trait and organisms may further evolve. Evolution may be convergent with similar traits evolving in multiple species or divergent with diverse traits evolving in multiple species that came from a common ancestor. We can observe evidence of evolution by means of DNA code and the fossil record, and also by the existence of homologous and vestigial structures.

Review Questions

Which scientific concept did Charles Darwin and Alfred Wallace independently discover?

1. mutation
 2. natural selection
 3. overbreeding
 4. sexual reproduction
-

B

Which of the following situations will lead to natural selection?

1. The seeds of two plants land near each other and one grows larger than the other.
 2. Two types of fish eat the same kind of food, and one is better able to gather food than the other.
 3. Male lions compete for the right to mate with females, with only one possible winner.
 4. all of the above
-

D

Which description is an example of a phenotype?

1. A certain duck has a blue beak.
2. A mutation occurred to a flower.
3. Most cheetahs live solitary lives.

4. both a and c

D

Which situation is most likely an example of convergent evolution?

1. Squid and humans have eyes similar in structure.
 2. Worms and snakes both move without legs.
 3. Some bats and birds have wings that allow them to fly.
 4. all of the above
-

D

Critical Thinking Questions

If a person scatters a handful of garden pea plant seeds in one area, how would natural selection work in this situation?

The plants that can best use the resources of the area, including competing with other individuals for those resources will produce

more seeds themselves and those traits that allowed them to better use the resources will increase in the population of the next generation.

Why do scientists consider vestigial structures evidence for evolution?

Vestigial structures are considered evidence for evolution because most structures do not exist in an organism without serving some function either presently or in the past. A vestigial structure indicates a past form or function that has since changed, but the structure remains present because it had a function in the ancestor.

How does the scientific meaning of “theory” differ from the common vernacular meaning?

In science, a theory is a thoroughly tested and verified set of explanations for a body of observations of nature. It is the strongest form of knowledge in science. In contrast, a theory in common vernacular can mean a guess or speculation about something, meaning that the knowledge implied by the theory is very weak.

Explain why the statement that a monkey is more evolved than a mouse is incorrect.

The statement implies that there is a goal to evolution and that the monkey represents greater progress to that goal than the mouse. Both species are likely to be well adapted to their particular environments, which is the outcome of natural selection.

Glossary

adaptation

heritable trait or behavior in an organism that aids in its survival and reproduction in its present environment

convergent evolution

process by which groups of organisms independently evolve to similar forms

divergent evolution

process by which groups of organisms evolve in diverse directions from a common point

homologous structures

parallel structures in diverse organisms that have a common ancestor

natural selection

reproduction of individuals with favorable genetic traits that survive environmental change because of those traits, leading to evolutionary change

variation

genetic differences among individuals in a population

vestigial structure

physical structure present in an organism but that has no apparent function and appears to be from a functional structure in a distant ancestor

Formation of New Species

By the end of this section, you will be able to do the following:

- Define species and describe how scientists identify species as different
- Describe genetic variables that lead to speciation
- Identify prezygotic and postzygotic reproductive barriers
- Explain allopatric and sympatric speciation
- Describe adaptive radiation

Although all life on earth shares various genetic similarities, only certain organisms combine genetic information by sexual reproduction and have offspring that can then successfully reproduce. Scientists call such organisms members of the same biological species.

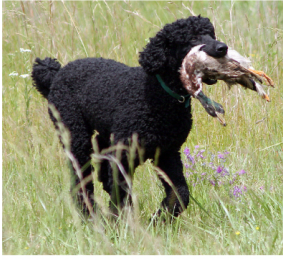
The (a) poodle and (b) cocker spaniel can reproduce to produce a breed known as (c) the cockapoo. (credit a: modification of work by Sally Eller, Tom Reese; credit b: modification of work by Jeremy McWilliams; credit c: modification of work by Kathleen Conklin) The (a) African fish eagle is similar in appearance to the (b) bald eagle, but the two birds are members of different species. (credit a: modification of work by Nigel Wedge; credit b: modification of work by U.S. Fish and Wildlife Service)

Species and the Ability to Reproduce

A **species** is a group of individual organisms that interbreed and produce fertile, viable offspring. According to this definition, one species is distinguished from another when, in nature, it is not possible for matings between individuals from each species to produce fertile offspring.

Members of the same species share both external and internal characteristics, which develop from their DNA. The closer relationship two organisms share, the more DNA they have in common, just like people and their families. People's DNA is likely to be more like their father or mother's DNA than their cousin or grandparent's DNA. Organisms of the same species have the highest level of DNA alignment and therefore share characteristics and behaviors that lead to successful reproduction.

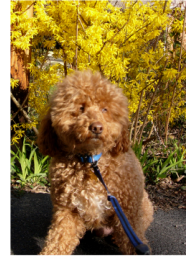
Species' appearance can be misleading in suggesting an ability or inability to mate. For example, even though domestic dogs (*Canis lupus familiaris*) display phenotypic differences, such as size, build, and coat, most dogs can interbreed and produce viable puppies that can mature and sexually reproduce ([link](#)).



(a)



(b)



(c)

In other cases, individuals may appear similar although they are not members of the same species. For example, even though bald eagles (*Haliaeetus leucocephalus*) and African fish eagles (*Haliaeetus vocifer*) are both birds and eagles, each belongs to a separate species group ([\[link\]](#)). If humans were to artificially intervene and fertilize a bald eagle's egg with an African fish eagle's sperm and a chick did hatch, that offspring, called a **hybrid** (a cross between two species), would probably be infertile—unable to successfully reproduce after it reached maturity. Different species may have different genes that are active in development; therefore, it may not be possible to develop a viable offspring with two different sets of directions. Thus, even though hybridization may take place, the two species still remain separate.



(a)



(b)

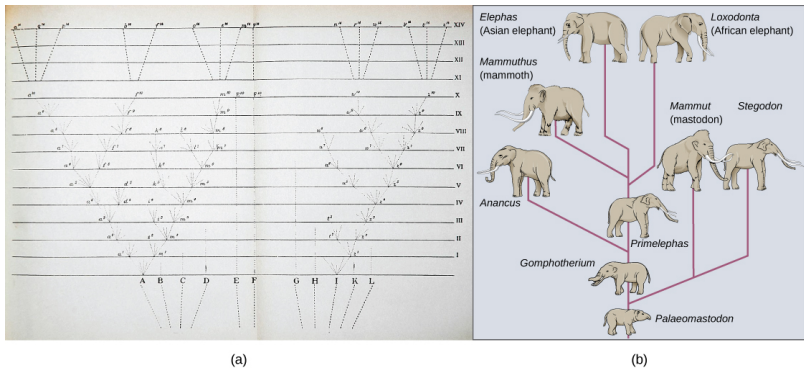
Populations of species share a gene pool: a collection of all the gene variants in the species. Again, the basis to any changes in a group or population of organisms must be genetic for this is the only way to share and pass on traits. When variations occur within a species, they can only pass to the next generation along two main pathways: asexual reproduction or sexual reproduction. The change will pass on asexually simply if the reproducing cell possesses the changed trait. For the changed trait to pass on by sexual reproduction, a gamete, such as a sperm or egg cell, must possess the changed trait. In other words, sexually-reproducing organisms can experience several genetic changes in their body cells, but if these changes do not occur in a sperm or egg cell, the changed trait will never reach the next generation. Only heritable traits can evolve. Therefore, reproduction plays a paramount role for genetic change to take root in a population or species. In short, organisms must be able to reproduce with each other to pass new traits to offspring.

The only illustration in Darwin's *On the Origin of Species* is (a) a diagram showing speciation events leading to biological diversity. The diagram shows similarities to phylogenetic charts that today illustrate the relationships of species. (b) Modern elephants evolved from the *Palaeomastodon*, a species that lived in Egypt 35–50 million years ago.

Speciation

The biological definition of species, which works for sexually reproducing organisms, is a group of actual or potential interbreeding individuals. There are exceptions to this rule. Many species are similar enough that hybrid offspring are possible and may often occur in nature, but for the majority of species this rule generally holds. The presence in nature of hybrids between similar species suggests that they may have descended from a single interbreeding species, and the speciation process may not yet be completed.

Given the extraordinary diversity of life on the planet there must be mechanisms for **speciation**: the formation of two species from one original species. Darwin envisioned this process as a branching event and diagrammed the process in the only illustration in *On the Origin of Species* ([\[link\]](#)^a). Compare this illustration to the diagram of elephant evolution ([\[link\]](#)), which shows that as one species changes over time, it branches to form more than one new species, repeatedly, as long as the population survives or until the organism becomes extinct.



For speciation to occur, two new populations must form from one original population and they must evolve in such a way that it becomes impossible for individuals from the two new populations to interbreed. Biologists have proposed mechanisms by which this could occur that fall into two broad categories. **Allopatric speciation** (allo- = "other"; -patric = "homeland") involves geographic separation of populations from a parent species and subsequent evolution. **Sympatric speciation** (sym- = "same"; -patric = "homeland") involves speciation occurring within a parent species remaining in one location.

Biologists think of speciation events as the splitting of one ancestral species into two descendant species. There is no reason why more than two species might not form at one time except that it is less likely and we can conceptualize multiple events as single splits occurring close in time.

The northern spotted owl and the Mexican spotted owl inhabit geographically separate locations with

different climates and ecosystems. The owl is an example of allopatric speciation. (credit "northern spotted owl": modification of work by John and Karen Hollingsworth; credit "Mexican spotted owl": modification of work by Bill Radke) The honeycreeper birds illustrate adaptive radiation. From one original species of bird, multiple others evolved, each with its own distinctive characteristics.

Allopatric Speciation

A geographically continuous population has a gene pool that is relatively homogeneous. Gene flow, the movement of alleles across a species' range, is relatively free because individuals can move and then mate with individuals in their new location. Thus, an allele's frequency at one end of a distribution will be similar to the allele's frequency at the other end. When populations become geographically discontinuous, it prevents alleles' free-flow. When that separation lasts for a period of time, the two populations are able to evolve along different trajectories. Thus, their allele frequencies at numerous genetic loci gradually become increasingly different as new alleles independently arise by mutation in each population. Typically, environmental conditions, such as climate, resources, predators, and competitors for the two populations will differ causing natural selection to favor divergent adaptations in each group.

Isolation of populations leading to allopatric speciation can occur in a variety of ways: a river forming a new branch, erosion creating a new valley, a group of organisms traveling to a new location without the ability to return, or seeds floating over the ocean to an island. The nature of the geographic separation necessary to isolate populations depends entirely on the organism's biology and its potential for dispersal. If two flying insect populations took up residence in separate nearby valleys, chances are, individuals from each population would fly back and forth continuing gene flow. However, if a new lake divided two rodent populations continued gene flow would be unlikely; therefore, speciation would be more likely.

Biologists group allopatric processes into two categories: dispersal and vicariance. **Dispersal** is when a few members of a species move to a new geographical area, and **vicariance** is when a natural situation arises to physically divide organisms.

Scientists have documented numerous cases of allopatric speciation taking place. For example, along the west coast of the United States, two separate spotted owl subspecies exist. The northern spotted owl has genetic and phenotypic differences from its close relative: the Mexican spotted owl, which lives in the south ([\[link\]](#)).

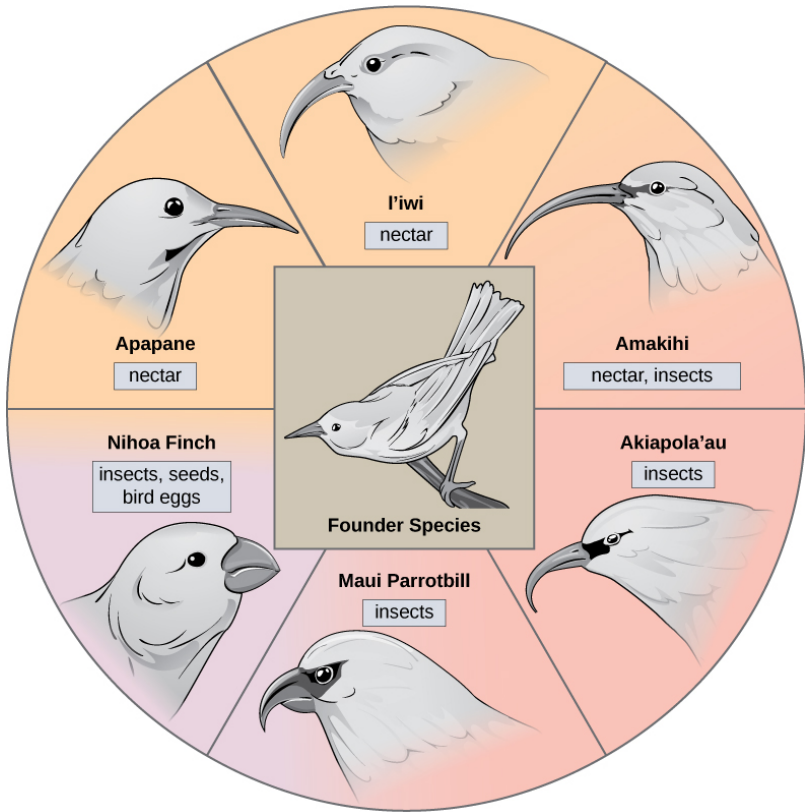


Additionally, scientists have found that the further the distance between two groups that once were the same species, the more likely it is that speciation will occur. This seems logical because as the distance increases, the various environmental factors would likely have less in common than locations in close proximity. Consider the two owls: in the north, the climate is cooler than in the south. The types of organisms in each ecosystem differ, as do their behaviors and habits. Also, the hunting habits and prey choices of the southern owls vary

from the northern owls. These variances can lead to evolved differences in the owls, and speciation likely will occur.

Adaptive Radiation

In some cases, a population of one species disperses throughout an area, and each finds a distinct niche or isolated habitat. Over time, the varied demands of their new lifestyles lead to multiple speciation events originating from a single species. We call this **adaptive radiation** because many adaptations evolve from a single point of origin; thus, causing the species to radiate into several new ones. Island archipelagos like the Hawaiian Islands provide an ideal context for adaptive radiation events because water surrounds each island which leads to geographical isolation for many organisms. The Hawaiian honeycreeper illustrates one example of adaptive radiation. From a single species, the founder species, numerous species have evolved, including the six in [\[link\]](#).



Notice the differences in the species' beaks in [\[link\]](#). Evolution in response to natural selection based on specific food sources in each new habitat led to evolution of a different beak suited to the specific food source. The seed-eating bird has a thicker, stronger beak which is suited to break hard nuts. The nectar-eating birds have long beaks to dip into flowers to reach the nectar. The insect-eating birds have beaks like swords, appropriate for stabbing and impaling insects. Darwin's finches are another example of adaptive radiation in an archipelago.

Link to Learning

Click through this [interactive site](#) to see how island birds evolved in evolutionary increments from 5 million years ago to today.

Autopolyploidy results when mitosis is not followed by cytokinesis. Allopolyploidy results when two species mate to produce viable offspring. In this example, a normal gamete from one species fuses with a polyploid gamete from another. Two matings are necessary to produce viable offspring.

Sympatric Speciation

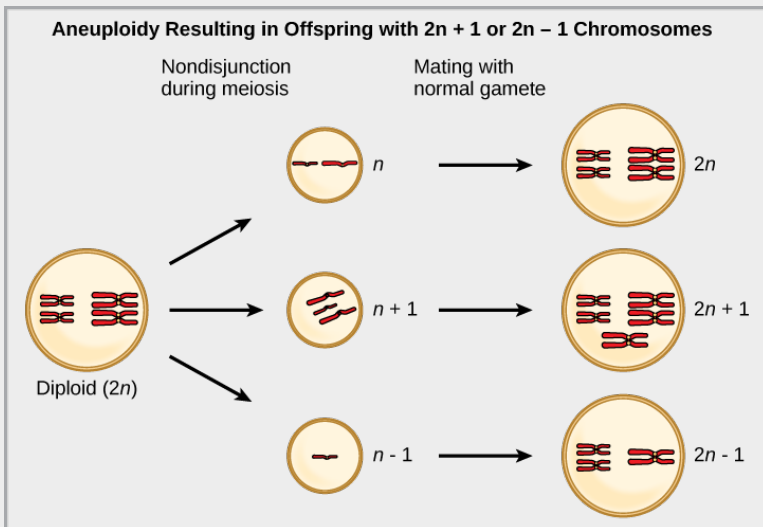
Can divergence occur if no physical barriers are in place to separate individuals who continue to live and reproduce in the same habitat? The answer is yes. We call the process of speciation within the same space sympatric. The prefix “sym” means same, so “sympatric” means “same homeland” in contrast to “allopatric” meaning “other homeland.” Scientists have proposed and studied many mechanisms.

One form of sympatric speciation can begin with a serious chromosomal error during cell division. In a normal cell division event chromosomes replicate, pair up, and then separate so that each new cell has the same number of chromosomes. However,

sometimes the pairs separate and the end cell product has too many or too few individual chromosomes in a condition that we call **aneuploidy** ([\[link\]](#)).

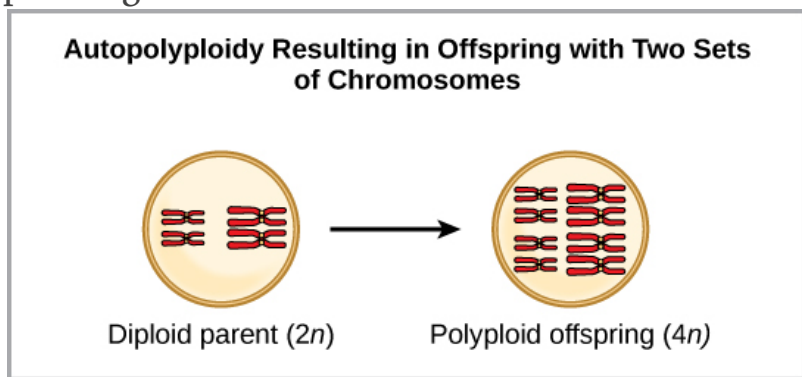
Visual Connection

Aneuploidy results when the gametes have too many or too few chromosomes due to nondisjunction during meiosis. In this example, the resulting offspring will have $2n + 1$ or $2n - 1$ chromosomes.



Which is most likely to survive, offspring with $2n + 1$ chromosomes or offspring with $2n - 1$ chromosomes?

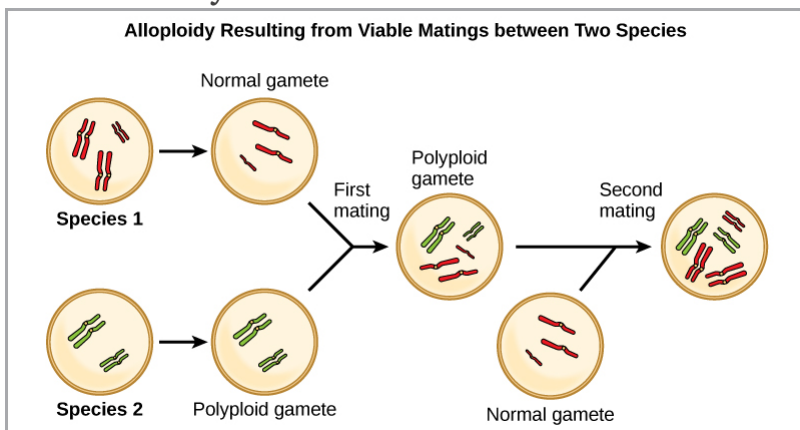
Polyploidy is a condition in which a cell or organism has an extra set, or sets, of chromosomes. Scientists have identified two main types of polyploidy that can lead to reproductive isolation of an individual in the polyploidy state. Reproductive isolation is the inability to interbreed. In some cases, a polyploid individual will have two or more complete sets of chromosomes from its own species in a condition that we call **autopolyploidy** ([\[link\]](#)). The prefix “auto-” means “self,” so the term means multiple chromosomes from one’s own species. Polyploidy results from an error in meiosis in which all of the chromosomes move into one cell instead of separating.



For example, if a plant species with $2n = 6$ produces autopolyploid gametes that are also diploid ($2n = 6$, when they should be $n = 3$), the gametes now have twice as many chromosomes as they should have. These new gametes will be incompatible with the normal gametes that this plant species produces. However, they could either self-pollinate or reproduce with other autopolyploid

plants with gametes having the same diploid number. In this way, sympatric speciation can occur quickly by forming offspring with $4n$ that we call a tetraploid. These individuals would immediately be able to reproduce only with those of this new kind and not those of the ancestral species.

The other form of polyploidy occurs when individuals of two different species reproduce to form a viable offspring that we call an **allopolyploid**. The prefix “allo-” means “other” (recall from allopatric): therefore, an allopolyploid occurs when gametes from two different species combine. [\[link\]](#) illustrates one possible way an allopolyploid can form. Notice how it takes two generations, or two reproductive acts, before the viable fertile hybrid results.



The cultivated forms of wheat, cotton, and tobacco plants are all allopolyploids. Although polyploidy occurs occasionally in animals, it takes place most commonly in plants. (Animals with any of the types

of chromosomal aberrations that we describe here are unlikely to survive and produce normal offspring.) Scientists have discovered more than half of all plant species studied relate back to a species evolved through polyploidy. With such a high rate of polyploidy in plants, some scientists hypothesize that this mechanism takes place more as an adaptation than as an error.

These two related frog species exhibit temporal reproductive isolation. (a) *Rana aurora* breeds earlier in the year than (b) *Rana boylei*. (credit a: modification of work by Mark R. Jennings, USFWS; credit b: modification of work by Alessandro Catenazzi) Speciation can occur when two populations occupy different habitats. The habitats need not be far apart. The cricket (a) *Gryllus pennsylvanicus* prefers sandy soil, and the cricket (b) *Gryllus firmus* prefers loamy soil. The two species can live in close proximity, but because of their different soil preferences, they became genetically isolated. The shape of the male reproductive organ varies among male damselfly species, and is only compatible with the female of that species. Reproductive organ incompatibility keeps the species reproductively isolated. Some flowers have evolved to attract certain pollinators. The (a) wide foxglove flower is adapted for pollination by bees, while the (b) long, tube-shaped trumpet creeper flower is adapted for pollination by hummingbirds. Cichlid fish from Lake Apoyeque, Nicaragua, show evidence of sympatric speciation. Lake Apoyeque, a

crater lake, is 1800 years old, but genetic evidence indicates that a single population of cichlid fish populated the lake only 100 years ago. Nevertheless, two populations with distinct morphologies and diets now exist in the lake, and scientists believe these populations may be in an early stage of speciation.

Reproductive Isolation

Given enough time, the genetic and phenotypic divergence between populations will affect characters that influence reproduction: if individuals of the two populations were brought together, mating would be less likely, but if mating occurred, offspring would be nonviable or infertile. Many types of diverging characters may affect the **reproductive isolation**, the ability to interbreed, of the two populations.

Reproductive isolation can take place in a variety of ways. Scientists organize them into two groups: prezygotic barriers and postzygotic barriers. Recall that a zygote is a fertilized egg: the first cell of an organism's development that reproduces sexually. Therefore, a **prezygotic barrier** is a mechanism that blocks reproduction from taking place. This includes barriers that prevent fertilization when organisms attempt reproduction. A **postzygotic barrier** occurs after zygote formation. This includes organisms that don't survive the embryonic stage and those that are

born sterile.

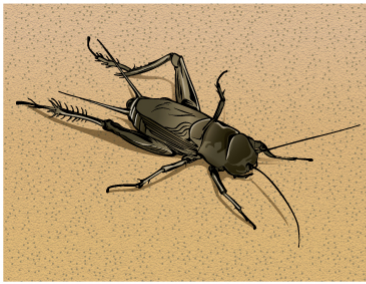
Some types of prezygotic barriers prevent reproduction entirely. Many organisms only reproduce at certain times of the year, often just annually. Differences in breeding schedules, which we call **temporal isolation**, can act as a form of reproductive isolation. For example, two frog species inhabit the same area, but one reproduces from January to March; whereas, the other reproduces from March to May ([\[link\]](#)).



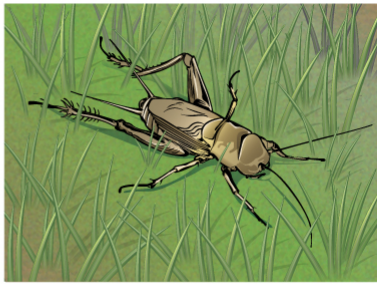
(a)

(b)

In some cases, populations of a species move or are moved to a new habitat and take up residence in a place that no longer overlaps with the same species' other populations. We call this situation **habitat isolation**. Reproduction with the parent species ceases, and a new group exists that is now reproductively and genetically independent. For example, a cricket population that was divided after a flood could no longer interact with each other. Over time, natural selection forces, mutation, and genetic drift will likely result in the two groups diverging ([\[link\]](#)).



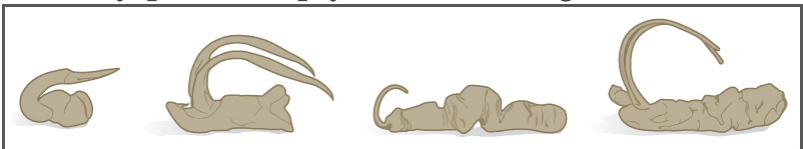
(a) *Gryllus pennsylvanicus* prefers sandy soil.



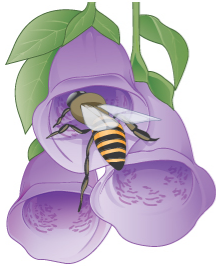
(b) *Gryllus firmus* prefers loamy soil.

Behavioral isolation occurs when the presence or absence of a specific behavior prevents reproduction. For example, male fireflies use specific light patterns to attract females. Various firefly species display their lights differently. If a male of one species tried to attract the female of another, she would not recognize the light pattern and would not mate with the male.

Other prezygotic barriers work when differences in their gamete cells (eggs and sperm) prevent fertilization from taking place. We call this a **gametic barrier**. Similarly, in some cases closely related organisms try to mate, but their reproductive structures simply do not fit together. For example, damselfly males of different species have differently shaped reproductive organs. If one species tries to mate with the female of another, their body parts simply do not fit together. ([link](#)).



In plants, certain structures aimed to attract one type of pollinator simultaneously prevent a different pollinator from accessing the pollen. The tunnel through which an animal must access nectar can vary widely in length and diameter, which prevents the plant from cross-pollinating with a different species ([\[link\]](#)).



(a) Honeybee drinking nectar from a foxglove flower



(b) Ruby-throated hummingbird drinking nectar from a trumpet creeper flower

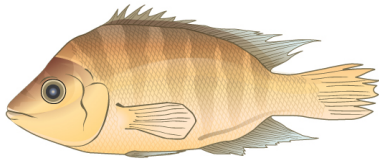
When fertilization takes place and a zygote forms, postzygotic barriers can prevent reproduction. Hybrid individuals in many cases cannot form normally in the womb and simply do not survive past the embryonic stages. We call this **hybrid inviability** because the hybrid organisms simply are not viable. In another postzygotic situation, reproduction leads to hybrid birth and growth that is sterile. Therefore, the organisms are unable to reproduce offspring of their own. We call this hybrid sterility.

Habitat Influence on Speciation

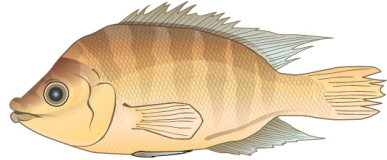
Sympatric speciation may also take place in ways other than polyploidy. For example, consider a fish

species that lives in a lake. As the population grows, competition for food increases. Under pressure to find food, suppose that a group of these fish had the genetic flexibility to discover and feed off another resource that other fish did not use. What if this new food source was located at a different depth of the lake? Over time, those feeding on the second food source would interact more with each other than the other fish; therefore, they would breed together as well. Offspring of these fish would likely behave as their parents: feeding and living in the same area and keeping separate from the original population. If this group of fish continued to remain separate from the first population, eventually sympatric speciation might occur as more genetic differences accumulated between them.

This scenario does play out in nature, as do others that lead to reproductive isolation. One such place is Lake Victoria in Africa, famous for its sympatric speciation of cichlid fish. Researchers have found hundreds of sympatric speciation events in these fish, which have not only happened in great number, but also over a short period of time. [\[link\]](#) shows this type of speciation among a cichlid fish population in Nicaragua. In this locale, two types of cichlids live in the same geographic location but have come to have different morphologies that allow them to eat various food sources.



Thin-lipped cichlid



Thick-lipped cichlid

Section Summary

Speciation occurs along two main pathways: geographic separation (allopatric speciation) and through mechanisms that occur within a shared habitat (sympatric speciation). Both pathways isolate a population reproductively in some form. Mechanisms of reproductive isolation act as barriers between closely related species, enabling them to diverge and exist as genetically independent species. Prezygotic barriers block reproduction prior to formation of a zygote; whereas, postzygotic barriers block reproduction after fertilization occurs. For a new species to develop, something must introduce a reproductive barrier. Sympatric speciation can occur through errors in meiosis that form gametes with extra chromosomes (polyploidy). Autopolyploidy occurs within a single species; whereas, allopolyploidy occurs between closely related species.

Visual Connection Questions

[link] Which is most likely to survive, offspring with $2n + 1$ chromosomes or offspring with $2n - 1$ chromosomes?

[link] Loss of genetic material is almost always lethal, so offspring with $2n + 1$ chromosomes are more likely to survive.

Review Questions

Which situation would most likely lead to allopatric speciation?

1. Flood causes the formation of a new lake.
2. A storm causes several large trees to fall down.
3. A mutation causes a new trait to develop.
4. An injury causes an organism to seek out a new food source.

A

What is the main difference between dispersal and vicariance?

1. One leads to allopatric speciation, whereas the other leads to sympatric speciation.
 2. One involves the movement of the organism, and the other involves a change in the environment.
 3. One depends on a genetic mutation occurring, and the other does not.
 4. One involves closely related organisms, and the other involves only individuals of the same species.
-

B

Which variable increases the likelihood of allopatric speciation taking place more quickly?

1. lower rate of mutation
 2. longer distance between divided groups
 3. increased instances of hybrid formation
 4. equivalent numbers of individuals in each population
-

B

What is the main difference between autopolyploid and allopolyploid?

1. the number of chromosomes

2. the functionality of the chromosomes
 3. the source of the extra chromosomes
 4. the number of mutations in the extra chromosomes
-

C

Which reproductive combination produces hybrids?

1. when individuals of the same species in different geographical areas reproduce
 2. when any two individuals sharing the same habitat reproduce
 3. when members of closely related species reproduce
 4. when offspring of the same parents reproduce
-

C

Which condition is the basis for a species to be reproductively isolated from other members?

1. It does not share its habitat with related species.
2. It does not exist out of a single habitat.
3. It does not exchange genetic information

with other species.

4. It does not undergo evolutionary changes for a significant period of time.

C

Which situation is *not* an example of a prezygotic barrier?

1. Two species of turtles breed at different times of the year.
2. Two species of flowers attract different pollinators.
3. Two species of birds display different mating dances.
4. Two species of insects produce infertile offspring.

D

Critical Thinking Questions

Why do island chains provide ideal conditions for adaptive radiation to occur?

Organisms of one species can arrive to an island together and then disperse throughout the chain, each settling into different niches and exploiting different food resources to reduce competition.

Two species of fish had recently undergone sympatric speciation. The males of each species had a different coloring through which the females could identify and choose a partner from her own species. After some time, pollution made the lake so cloudy that it was hard for females to distinguish colors. What might take place in this situation?

It is likely the two species would start to reproduce with each other. Depending on the viability of their offspring, they may fuse back into one species.

Why can polyploidy individuals lead to speciation fairly quickly?

The formation of gametes with new n numbers can occur in one generation. After a couple of generations, enough of these new hybrids can form to reproduce together as a new species.

Glossary

adaptive radiation

speciation when one species radiates to form several other species

allopatric speciation

speciation that occurs via geographic separation

allopolyploid

polyploidy formed between two related, but separate species

aneuploidy

condition of a cell having an extra chromosome or missing a chromosome for its species

autopolyploid

polyploidy formed within a single species

behavioral isolation

type of reproductive isolation that occurs when a specific behavior or lack of one prevents reproduction from taking place

dispersal

allopatric speciation that occurs when a few members of a species move to a new geographical area

gametic barrier

prezygotic barrier occurring when closely related individuals of different species mate, but differences in their gamete cells (eggs and sperm) prevent fertilization from taking place

habitat isolation

reproductive isolation resulting when species' populations move or are moved to a new habitat, taking up residence in a place that no longer overlaps with the same species' other populations

hybrid

offspring of two closely related individuals, not of the same species

postzygotic barrier

reproductive isolation mechanism that occurs after zygote formation

prezygotic barrier

reproductive isolation mechanism that occurs before zygote formation

reproductive isolation

situation that occurs when a species is reproductively independent from other species; behavior, location, or reproductive barriers may cause this to happen

speciation

formation of a new species

species

group of populations that interbreed and produce fertile offspring

sympatric speciation

speciation that occurs in the same geographic space

temporal isolation

differences in breeding schedules that can act as a form of prezygotic barrier leading to reproductive isolation

vicariance

allopatric speciation that occurs when something in the environment separates organisms of the same species into separate groups

Reconnection and Speciation Rates

By the end of this section, you will be able to do the following:

- Describe pathways of species evolution in hybrid zones
- Explain the two major theories on rates of speciation

Speciation occurs over a span of evolutionary time, so when a new species arises, there is a transition period during which the closely related species continue to interact.

Reconnection

After speciation, two species may recombine or even continue interacting indefinitely. Individual organisms will mate with any nearby individual with whom they are capable of breeding. We call an area where two closely related species continue to interact and reproduce, forming hybrids a **hybrid zone**. Over time, the hybrid zone may change depending on the fitness of the hybrids and the reproductive barriers ([\[link\]](#)). If the hybrids are less fit than the parents, speciation reinforcement occurs, and the species continue to diverge until they can no longer mate and produce viable offspring. If reproductive barriers weaken, fusion

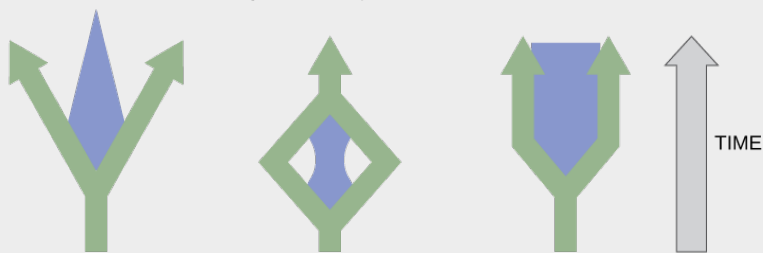
occurs and the two species become one. Barriers remain the same if hybrids are fit and reproductive: stability may occur and hybridization continues.

Visual Connection

After speciation has occurred, the two separate but closely related species may continue to produce offspring in an area called the hybrid zone.

Reinforcement, fusion, or stability may result, depending on reproductive barriers and the relative fitness of the hybrids.

Changes in the Hybrid Zone over Time



Reinforcement:

Hybrids are less fit than either purebred species. The species continue to diverge until hybridization can no longer occur.

Fusion:

Reproductive barriers weaken until the two species become one.

Stability:

Fit hybrids continue to be produced.

If two species eat a different diet but one of the food sources is eliminated and both species are forced to eat the same foods, what change in the hybrid zone is most likely to occur?

Hybrids can be either less fit than the parents, more

fit, or about the same. Usually hybrids tend to be less fit; therefore, such reproduction diminishes over time, nudging the two species to diverge further in a process we call **reinforcement**. Scientists use this term because the hybrids' low success reinforces the original speciation. If the hybrids are as fit or more fit than the parents, the two species may fuse back into one species ([\[link\]](#)). Scientists have also observed that sometimes two species will remain separate but also continue to interact to produce some individuals. Scientists classify this as stability because no real net change is taking place.

Varying Rates of Speciation

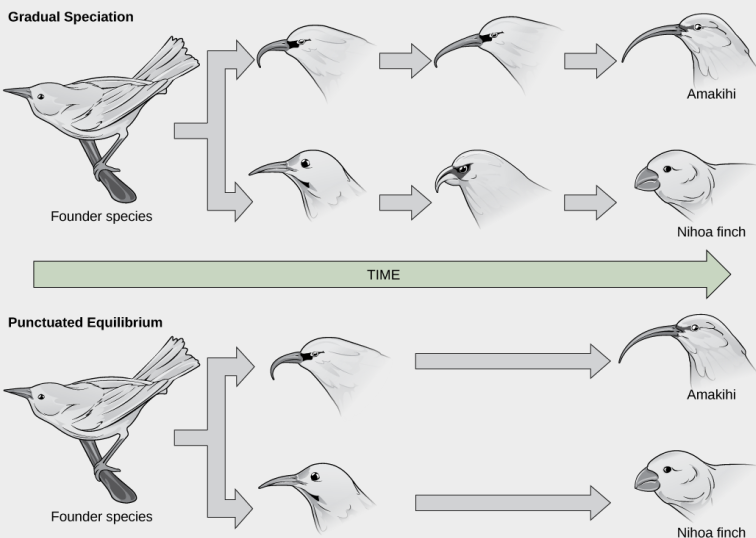
Scientists around the world study speciation, documenting observations both of living organisms and those found in the fossil record. As their ideas take shape and as research reveals new details about how life evolves, they develop models to help explain speciation rates. In terms of how quickly speciation occurs, we can observe two current patterns: gradual speciation model and punctuated equilibrium model.

In the **gradual speciation model**, species diverge gradually over time in small steps. In the **punctuated equilibrium** model, a new species undergoes changes quickly from the parent species, and then remains largely unchanged for long

periods of time afterward ([\[link\]](#)). We call this early change model punctuated equilibrium, because it begins with a punctuated or periodic change and then remains in balance afterward. While punctuated equilibrium suggests a faster tempo, it does not necessarily exclude gradualism.

Visual Connection

In (a) gradual speciation, species diverge at a slow, steady pace as traits change incrementally. In (b) punctuated equilibrium, species diverge quickly and then remain unchanged for long periods of time.



Which of the following statements is false?

1. Punctuated equilibrium is most likely to occur in a small population that experiences a rapid

change in its environment.

2. Punctuated equilibrium is most likely to occur in a large population that lives in a stable climate.
3. Gradual speciation is most likely to occur in species that live in a stable climate.
4. Gradual speciation and punctuated equilibrium both result in the divergence of species.

The primary influencing factor on changes in speciation rate is environmental conditions. Under some conditions, selection occurs quickly or radically. Consider a species of snails that had been living with the same basic form for many thousands of years. Layers of their fossils would appear similar for a long time. When a change in the environment takes place—such as a drop in the water level—a small number of organisms are separated from the rest in a brief period of time, essentially forming one large and one tiny population. The tiny population faces new environmental conditions. Because its gene pool quickly became so small, any variation that surfaces and that aids in surviving the new conditions becomes the predominant form.

Link to Learning

Visit [this website](#) to continue the speciation story of the snails.

Section Summary

Speciation is not a precise division: overlap between closely related species can occur in areas called hybrid zones. Organisms reproduce with other similar organisms. The fitness of these hybrid offspring can affect the two species' evolutionary path. Scientists propose two models for the rate of speciation: one model illustrates how a species can change slowly over time. The other model demonstrates how change can occur quickly from a parent generation to a new species. Both models continue to follow natural selection patterns.

Visual Connection Questions

[\[link\]](#) If two species eat a different diet but one of the food sources is eliminated and both species are forced to eat the same foods, what change in the hybrid zone is most likely to occur?

[\[link\]](#) Fusion is most likely to occur because the two species will interact more and similar traits in food acquisition will be selected.

[\[link\]](#) Which of the following statements is false?

1. Punctuated equilibrium is most likely to occur in a small population that experiences a rapid change in its environment.
2. Punctuated equilibrium is most likely to occur in a large population that lives in a stable climate.
3. Gradual speciation is most likely to occur in species that live in a stable climate.
4. Gradual speciation and punctuated equilibrium both result in the evolution of new species.

[\[link\]](#) Answer B

Review Questions

Which term is used to describe the continued divergence of species based on the low fitness

of hybrid offspring?

1. reinforcement
2. fusion
3. stability
4. punctuated equilibrium

A

Which components of speciation would be least likely to be a part of punctuated equilibrium?

1. a division of populations
2. a change in environmental conditions
3. ongoing gene flow among all individuals
4. a large number of mutations taking place at once

C

Critical Thinking Questions

What do both rate of speciation models have in common?

Both models continue to conform to the rules of natural selection, and the influences of gene flow, genetic drift, and mutation.

Describe a situation where hybrid reproduction would cause two species to fuse into one.

If the hybrid offspring are as fit or more fit than the parents, reproduction would likely continue between both species and the hybrids, eventually bringing all organisms under the umbrella of one species.

Glossary

gradual speciation model

model that shows how species diverge gradually over time in small steps

hybrid zone

area where two closely related species continue to interact and reproduce, forming hybrids

punctuated equilibrium

model for rapid speciation that can occur when an event causes a small portion of a population to be cut off from the rest of the population

reinforcement

continued speciation divergence between two
related species due to low fitness of hybrids
between them

Introduction

class = "introduction" Living things may be single-celled or complex, multicellular organisms. They may be plants, animals, fungi, bacteria, or archaea. This diversity results from evolution. (credit "wolf": modification of work by Gary Kramer; credit "coral": modification of work by William Harrigan, NOAA; credit "river": modification of work by Vojtěch Dostál; credit "fish" modification of work by Christian Mehlführer; credit "mushroom": modification of work by Cory Zanker; credit "tree": modification of work by Joseph Kranak; credit "bee": modification of work by Cory Zanker)



All life on Earth is related. Evolutionary theory states that humans, beetles, plants, and bacteria all share a common ancestor, but that millions of years of evolution have shaped each of these organisms into the forms we see today. Scientists consider evolution a key concept to understanding life. It is one of the most dominant evolutionary forces.

Natural selection acts to promote traits and behaviors that increase an organism's chances of survival and reproduction, while eliminating those traits and behaviors that are detrimental to the organism. However, natural selection can only, as its name implies, select—it cannot create. We can attribute novel traits and behaviors to another evolutionary force—mutation. Mutation and other sources of variation among individuals, as well as the evolutionary forces that act upon them, alter populations and species. This combination of processes has led to the world of life we see today.

Population Evolution

By the end of this section, you will be able to do the following:

- Define population genetics and describe how scientists use population genetics in studying population evolution
- Define the Hardy-Weinberg principle and discuss its importance

People did not understand the mechanisms of inheritance, or genetics, at the time Charles Darwin and Alfred Russel Wallace were developing their idea of natural selection. This lack of knowledge was a stumbling block to understanding many aspects of evolution. The predominant (and incorrect) genetic theory of the time, blending inheritance, made it difficult to understand how natural selection might operate. Darwin and Wallace were unaware of the Austrian monk Gregor Mendel's 1866 publication "Experiments in Plant Hybridization", which came out not long after Darwin's book, *On the Origin of Species*. Scholars rediscovered Mendel's work in the early twentieth century at which time geneticists were rapidly coming to an understanding of the basics of inheritance. Initially, the newly discovered particulate nature of genes made it difficult for biologists to understand how gradual evolution could occur. However, over the next few decades scientists integrated genetics and evolution in what

became known as the **modern synthesis**—the coherent understanding of the relationship between natural selection and genetics that took shape by the 1940s. Generally, this concept is generally accepted today. In short, the modern synthesis describes how evolutionary processes, such as natural selection, can affect a population's genetic makeup, and, in turn, how this can result in the gradual evolution of populations and species. The theory also connects population change over time (**microevolution**), with the processes that gave rise to new species and higher taxonomic groups with widely divergent characters, called (**macroevolution**).

Everyday Connection

Evolution and Flu Vaccines

Every fall, the media starts reporting on flu vaccinations and potential outbreaks. Scientists, health experts, and institutions determine recommendations for different parts of the population, predict optimal production and inoculation schedules, create vaccines, and set up clinics to provide inoculations. You may think of the annual flu shot as media hype, an important health protection, or just a briefly uncomfortable prick in your arm. However, do you think of it in terms of evolution?

The media hype of annual flu shots is scientifically grounded in our understanding of evolution. Each

year, scientists across the globe strive to predict the flu strains that they anticipate as most widespread and harmful in the coming year. They base this knowledge on how flu strains have evolved over time and over the past few flu seasons. Scientists then work to create the most effective vaccine to combat those selected strains. Pharmaceutical companies produce hundreds of millions of doses in a short period in order to provide vaccinations to key populations at the optimal time.

Because viruses, like the flu, evolve very quickly (especially in evolutionary time), this poses quite a challenge. Viruses mutate and replicate at a fast rate, so the vaccine developed to protect against last year's flu strain may not provide the protection one needs against the coming year's strain.

Evolution of these viruses means continued adaptations to ensure survival, including adaptations to survive previous vaccines.

Sahar S. Hanania, Dhia S. Hassawi, and Nidal M. Irshaid, "Allele Frequency and Molecular Genotypes of ABO Blood Group System in a Jordanian Population," *Journal of Medical Sciences* 7 (2007): 51-58, doi:10.3923/jms.2007.51.58.

Population Genetics

Recall that a gene for a particular character may

have several alleles, or variants, that code for different traits associated with that character. For example, in the ABO blood type system in humans, three alleles determine the particular blood-type carbohydrate on the surface of red blood cells. Each individual in a population of diploid organisms can only carry two alleles for a particular gene, but more than two may be present in the individuals that comprise the population. Mendel followed alleles as they were inherited from parent to offspring. In the early twentieth century, biologists in the area of **population genetics** began to study how selective forces change a population through changes in allele and genotypic frequencies.

The **allele frequency** (or gene frequency) is the rate at which a specific allele appears within a population. Until now we have discussed evolution as a change in the characteristics of a population of organisms, but behind that phenotypic change is genetic change. In population genetics, scientists define the term evolution as a change in the allele's frequency in a population. Using the ABO blood type system as an example, the frequency of one of the alleles, I_A , is the number of copies of that allele divided by all the copies of the ABO gene in the population. For example, a study in Jordan^[footnote] found a frequency of I_A to be 26.1 percent. The I_B and I_O alleles comprise 13.4 percent and 60.5 percent of the alleles respectively, and all of the frequencies added up to 100 percent. A change in

this frequency over time would constitute evolution in the population.

The allele frequency within a given population can change depending on environmental factors; therefore, certain alleles become more widespread than others during the natural selection process. Natural selection can alter the population's genetic makeup. An example is if a given allele confers a phenotype that allows an individual to better survive or have more offspring. Because many of those offspring will also carry the beneficial allele, and often the corresponding phenotype, they will have more offspring of their own that also carry the allele, thus, perpetuating the cycle. Over time, the allele will spread throughout the population. Some alleles will quickly become fixed in this way, meaning that every individual of the population will carry the allele, while detrimental mutations may be swiftly eliminated if derived from a dominant allele from the gene pool. The **gene pool** is the sum of all the alleles in a population.

Sometimes, allele frequencies within a population change randomly with no advantage to the population over existing allele frequencies. We call this phenomenon genetic drift. Natural selection and genetic drift usually occur simultaneously in populations and are not isolated events. It is hard to determine which process dominates because it is often nearly impossible to determine the cause of

change in allele frequencies at each occurrence. We call an event that initiates an allele frequency change in an isolated part of the population, which is not typical of the original population, the **founder effect**. Natural selection, random drift, and founder effects can lead to significant changes in a population's genome.

Hardy-Weinberg Principle of Equilibrium

In the early twentieth century, English mathematician Godfrey Hardy and German physician Wilhelm Weinberg stated the principle of equilibrium to describe the population's genetic makeup. The theory, which later became known as the Hardy-Weinberg principle of equilibrium, states that a population's allele and genotype frequencies are inherently stable— unless some kind of evolutionary force is acting upon the population, neither the allele nor the genotypic frequencies would change. The Hardy-Weinberg principle assumes conditions with no mutations, migration, emigration, or selective pressure for or against genotype, plus an infinite population. While no population can satisfy those conditions, the principle offers a useful model against which to compare real population changes.

Working under this theory, population geneticists represent different alleles as different variables in

their mathematical models. The variable p , for example, often represents the frequency of a particular allele, say Y for the trait of yellow in Mendel's peas, while the variable q represents the frequency of y alleles that confer the color green. If these are the only two possible alleles for a given locus in the population, $p + q = 1$. In other words, all the p alleles and all the q alleles comprise all of the alleles for that locus in the population.

However, what ultimately interests most biologists is not the frequencies of different alleles, but the frequencies of the resulting genotypes, known as the population's **genetic structure**, from which scientists can surmise phenotype distribution. If we observe the phenotype, we can know only the homozygous recessive allele's genotype. The calculations provide an estimate of the remaining genotypes. Since each individual carries two alleles per gene, if we know the allele frequencies (p and q), predicting the genotypes' frequencies is a simple mathematical calculation to determine the probability of obtaining these genotypes if we draw two alleles at random from the gene pool. In the above scenario, an individual pea plant could be pp (YY), and thus produce yellow peas; pq (Yy), also yellow; or qq (yy), and thus produce green peas ([\[link\]](#)). In other words, the frequency of pp individuals is simply p^2 ; the frequency of pq individuals is $2pq$; and the frequency of qq individuals is q^2 . Again, if p and q are the only two

possible alleles for a given trait in the population, these genotypes frequencies will sum to one: $p^2 + 2pq + q^2 = 1$.

Visual Connection

When populations are in the Hardy-Weinberg equilibrium, the allelic frequency is stable from generation to generation and we can determine the allele distribution from the Hardy-Weinberg equation. If the allelic frequency measured in the field differs from the predicted value, scientists can make inferences about what evolutionary forces are at play.

Hardy-Weinberg Principle

Parent generation

Phenotype

Genotypic frequency

Number of individuals
(total = 500)

Number of alleles
in gene pool
(total = 1000)

Allelic frequency

YY

.49

245

Yy

.42

210

yy

.09

45

Y: $490 + 210 = 700$

y: $210 + 90 = 300$

$\frac{700 \text{ Y}}{1000 \text{ total}} = .7 = p$

$\frac{300 \text{ y}}{1000 \text{ total}} = .3 = q$

Hardy-Weinberg analysis

p (.7)

q (.3)

p (.7)

YY

$p^2 = .49$

Yy

$pq = .21$

q (.3)

Yy

$pq = .21$

yy

$q^2 = .09$

p^2

+

$2pq$

+

q^2

= 1

.7²

+

$2(.7)(.3)$

+

.3²

= 1

.49

+

.42

+

.09

= 1

Predicted
frequency
of YY
offspring

Predicted
frequency
of Yy
offspring

Predicted
frequency
of yy
offspring

In plants, violet flower color (V) is dominant over white (v). If $p = 0.8$ and $q = 0.2$ in a population of 500 plants, how many individuals would you expect to be homozygous dominant (VV), heterozygous (Vv), and homozygous recessive (vv)? How many plants would you expect to have violet flowers, and how many would have white flowers?

In theory, if a population is at equilibrium—that is, there are no evolutionary forces acting upon it—generation after generation would have the same gene pool and genetic structure, and these equations would all hold true all of the time. Of course, even Hardy and Weinberg recognized that no natural population is immune to evolution. Populations in nature are constantly changing in genetic makeup due to drift, mutation, possibly migration, and selection. As a result, the only way to determine the exact distribution of phenotypes in a population is to go out and count them. However, the Hardy-Weinberg principle gives scientists a mathematical baseline of a non-evolving population to which they can compare evolving populations and thereby infer what evolutionary forces might be at play. If the frequencies of alleles or genotypes deviate from the value expected from the Hardy-Weinberg equation, then the population is evolving.

Link to Learning

Use this [online calculator](#) to determine a population's genetic structure.

Section Summary

The modern synthesis of evolutionary theory grew out of the cohesion of Darwin's, Wallace's, and Mendel's thoughts on evolution and heredity, along with the more modern study of population genetics. It describes the evolution of populations and species, from small-scale changes among individuals to large-scale changes over paleontological time periods. To understand how organisms evolve, scientists can track populations' allele frequencies over time. If they differ from generation to generation, scientists can conclude that the population is not in Hardy-Weinberg equilibrium, and is thus evolving.

Visual Connection Questions

[\[link\]](#) In plants, violet flower color (V) is dominant over white (v). If $p = .8$ and $q = 0.2$ in a population of 500 plants, how many individuals would you expect to be homozygous dominant (VV), heterozygous (Vv), and homozygous recessive (vv)? How many plants would you expect to have violet flowers, and how many would have white flowers?

[\[link\]](#) The expected distribution is 320 VV, 160Vv, and 20 vv plants. Plants with VV or Vv

genotypes would have violet flowers, and plants with the vv genotype would have white flowers, so a total of 480 plants would be expected to have violet flowers, and 20 plants would have white flowers.

Review Questions

What is the difference between micro- and macroevolution?

1. Microevolution describes the evolution of small organisms, such as insects, while macroevolution describes the evolution of large organisms, like people and elephants.
 2. Microevolution describes the evolution of microscopic entities, such as molecules and proteins, while macroevolution describes the evolution of whole organisms.
 3. Microevolution describes the evolution of organisms in populations, while macroevolution describes the evolution of species over long periods of time.
 4. Microevolution describes the evolution of organisms over their lifetimes, while macroevolution describes the evolution of organisms over multiple generations.
-

C

Population genetics is the study of:

1. how selective forces change the allele frequencies in a population over time
2. the genetic basis of population-wide traits
3. whether traits have a genetic basis
4. the degree of inbreeding in a population

A

Which of the following populations is not in Hardy-Weinberg equilibrium?

1. a population with 12 homozygous recessive individuals (yy), 8 homozygous dominant individuals (YY), and 4 heterozygous individuals (Yy)
2. a population in which the allele frequencies do not change over time
3. $p^2 + 2pq + q^2 = 1$
4. a population undergoing natural selection

D

One of the original Amish colonies rose from a

ship of colonists that came from Europe. The ship's captain, who had polydactyly, a rare dominant trait, was one of the original colonists. Today, we see a much higher frequency of polydactyly in the Amish population. This is an example of:

1. natural selection
2. genetic drift
3. founder effect
4. b and c

D

Critical Thinking Questions

Solve for the genetic structure of a population with 12 homozygous recessive individuals (yy), 8 homozygous dominant individuals (YY), and 4 heterozygous individuals (Yy).

$$p = (8 \cdot 2 + 4) / 48 = .42; q = (12 \cdot 2 + 4) / 48 = .58; p^2 = .17; 2pq = .48; q^2 = .34$$

Explain the Hardy-Weinberg principle of

equilibrium theory.

The Hardy-Weinberg principle of equilibrium is used to describe the genetic makeup of a population. The theory states that a population's allele and genotype frequencies are inherently stable: unless some kind of evolutionary force is acting upon the population, generation after generation of the population would carry the same genes, and individuals would, as a whole, look essentially the same.

Imagine you are trying to test whether a population of flowers is undergoing evolution. You suspect there is selection pressure on the color of the flower: bees seem to cluster around the red flowers more often than the blue flowers. In a separate experiment, you discover blue flower color is dominant to red flower color. In a field, you count 600 blue flowers and 200 red flowers. What would you expect the genetic structure of the flowers to be?

Red is recessive so $q^2 = 200/800 = 0.25$; $q = 0.5$; $p = 1 - q = 0.5$; $p^2 = 0.25$; $2pq = 0.5$. You would expect 200 homozygous blue flowers, 400 heterozygous blue flowers, and 200 red flowers.

Glossary

allele frequency

(also, gene frequency) rate at which a specific allele appears within a population

founder effect

event that initiates an allele frequency change in part of the population, which is not typical of the original population

gene pool

all the alleles that the individuals in the population carry

genetic structure

distribution of the different possible genotypes in a population

macroevolution

broader scale evolutionary changes that scientists see over paleontological time

microevolution

changes in a population's genetic structure

modern synthesis

overarching evolutionary paradigm that took shape by the 1940s and scientists generally accept today

population genetics

study of how selective forces change the allele frequencies in a population over time

Population Genetics

By the end of this section, you will be able to do the following:

- Describe the different types of variation in a population
- Explain why only natural selection can act upon heritable variation
- Describe genetic drift and the bottleneck effect
- Explain how each evolutionary force can influence a population's allele frequencies

A population's individuals often display different phenotypes, or express different alleles of a particular gene, which scientists refer to as polymorphisms. We call populations with two or more variations of particular characteristics polymorphic. A number of factors, including the population's genetic structure and the environment ([\[link\]](#)) influence **population variation**, the distribution of phenotypes among individuals. Understanding phenotypic variation sources in a population is important for determining how a population will evolve in response to different evolutionary pressures.

The distribution of phenotypes in this litter of kittens illustrates population variation. (credit: Pieter Lanser)



Genetic Variance

Natural selection and some of the other evolutionary forces can only act on heritable traits, namely an organism's genetic code. Because alleles are passed from parent to offspring, those that confer beneficial traits or behaviors may be selected, while deleterious alleles may not. Acquired traits, for the most part, are not heritable. For example, if an athlete works out in the gym every day, building up muscle strength, the athlete's offspring will not necessarily grow up to be a body builder. If there is a genetic basis for the ability to run fast, on the other hand, a parent may pass this to a child.

Link to Learning

Before Darwinian evolution became the prevailing theory of the field, French naturalist Jean-Baptiste Lamarck theorized that organisms could inherit acquired traits. While the majority of scientists have not supported this hypothesis, some have recently begun to realize that Lamarck was not completely wrong. Visit this [site](#) to learn more.

Heritability is the fraction of phenotype variation that we can attribute to genetic differences, or genetic variance, among individuals in a population. The greater the heritability of a population's phenotypic variation, the more susceptible it is to the evolutionary forces that act on heritable variation.

We call the diversity of alleles and genotypes within a population **genetic variance**. When scientists are involved in the breeding of a species, such as with animals in zoos and nature preserves, they try to increase a population's genetic variance to preserve as much of the phenotypic diversity as possible. This also helps reduce associated risks of **inbreeding**, the mating of closely related individuals, which can have the undesirable effect of bringing together deleterious recessive mutations that can cause abnormalities and susceptibility to disease. For example, a disease that is caused by a rare, recessive

allele might exist in a population, but it will only manifest itself when an individual carries two copies of the allele. Because the allele is rare in a normal, healthy population with unrestricted habitat, the chance that two carriers will mate is low, and even then, only 25 percent of their offspring will inherit the disease allele from both parents. While it is likely to happen at some point, it will not happen frequently enough for natural selection to be able to swiftly eliminate the allele from the population, and as a result, the allele maintains itself at low levels in the gene pool. However, if a family of carriers begins to interbreed with each other, this will dramatically increase the likelihood of two carriers mating and eventually producing diseased offspring, a phenomenon that scientists call **inbreeding depression**.

Changes in allele frequencies that we identify in a population can shed light on how it is evolving. In addition to natural selection, there are other evolutionary forces that could be in play: genetic drift, gene flow, mutation, nonrandom mating, and environmental variances.

A chance event or catastrophe can reduce the genetic variability within a population. A. J. Tipping et al., "Molecular and Genealogical Evidence for a Founder Effect in Fanconi Anemia Families of the Afrikaner Population of South Africa," *PNAS* 98, no. 10 (2001): 5734-5739, doi: 10.1073/pnas.091402398.

Genetic Drift

The theory of natural selection stems from the observation that some individuals in a population are more likely to survive longer and have more offspring than others; thus, they will pass on more of their genes to the next generation. A big, powerful male gorilla, for example, is much more likely than a smaller, weaker one to become the population's silverback, the pack's leader who mates far more than the other males of the group. The pack leader will father more offspring, who share half of his genes, and are likely to also grow bigger and stronger like their father. Over time, the genes for bigger size will increase in frequency in the population, and the population will, as a result, grow larger on average. That is, this would occur if this particular **selection pressure**, or driving selective force, were the only one acting on the population. In other examples, better camouflage or a stronger resistance to drought might pose a selection pressure.

Another way a population's allele and genotype frequencies can change is **genetic drift** ([\[link\]](#)), which is simply the effect of chance. By chance, some individuals will have more offspring than others—not due to an advantage conferred by some genetically-encoded trait, but just because one male happened to be in the right place at the right time (when the receptive female walked by) or because

the other one happened to be in the wrong place at the wrong time (when a fox was hunting).

Visual Connection

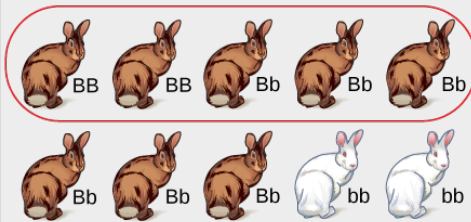
Genetic drift in a population can lead to eliminating an allele from a population by chance. In this example, rabbits with the brown coat color allele (B) are dominant over rabbits with the white coat color allele (b). In the first generation, the two alleles occur with equal frequency in the population, resulting in p and q values of .5. Only half of the individuals reproduce, resulting in a second generation with p and q values of .7 and .3, respectively. Only two individuals in the second generation reproduce, and by chance these individuals are homozygous dominant for brown coat color. As a result, in the third generation the recessive b allele is lost.

Genetic Drift

First generation

p (B gene frequency) = .5

q (b gene frequency) = .5



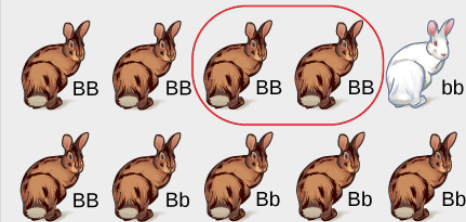
5 rabbits reproduce



Second generation

p = .7

q = .3



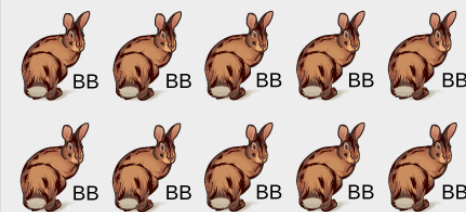
2 rabbits reproduce



Third generation

p = 1

q = 0



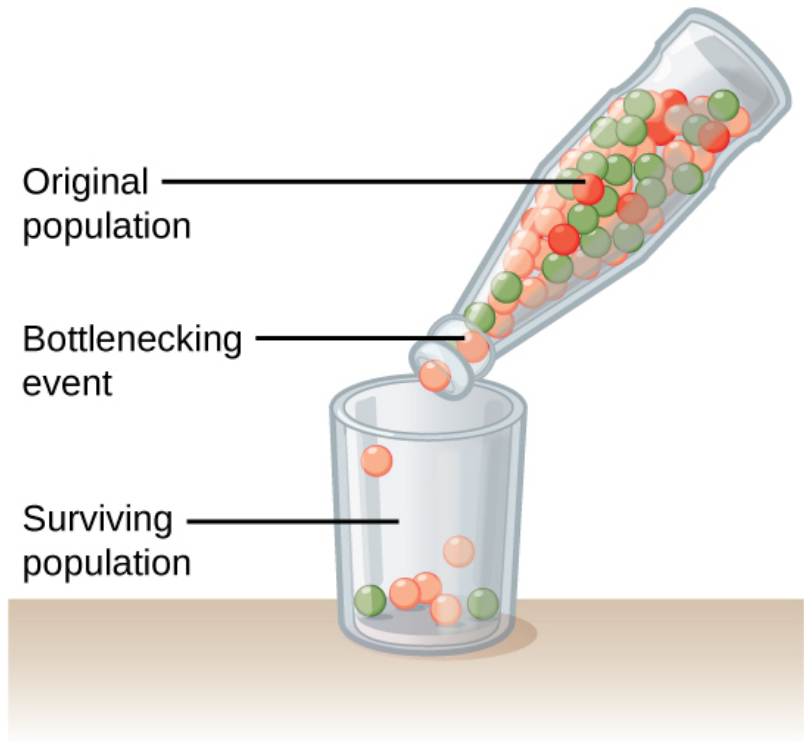
Do you think genetic drift would happen more quickly on an island or on the mainland?

Small populations are more susceptible to the forces of genetic drift. Large populations, alternatively, are buffered against the effects of chance. If one individual of a population of 10 individuals happens to die at a young age before it leaves any offspring to the next generation, all of its genes—1/10 of the population's gene pool—will be suddenly lost. In a population of 100, that's only 1 percent of the overall gene pool; therefore, it is much less impactful on the population's genetic structure.

Link to Learning

Go to this [site](#) to watch an animation of random sampling and genetic drift in action.

Natural events, such as an earthquake disaster that kills—at random—a large portion of the population, can magnify genetic drift. Known as the **bottleneck effect**, it results in suddenly wiping out a large portion of the genome ([\[link\]](#)). At once, the survivors' genetic structure becomes the entire population's genetic structure, which may be very different from the pre-disaster population.



Another scenario in which populations might experience a strong influence of genetic drift is if some portion of the population leaves to start a new population in a new location or if a physical barrier divides a population. In this situation, those individuals are an unlikely representation of the entire population, which results in the founder effect. The founder effect occurs when the genetic structure changes to match that of the new population's founding fathers and mothers. Researchers believe that the founder effect was a key factor in the genetic history of the Afrikaner population of Dutch settlers in South Africa, as evidenced by mutations that are common in

Afrikaners but rare in most other populations. This is probably because a higher-than-normal proportion of the founding colonists carried these mutations. As a result, the population expresses unusually high incidences of Huntington's disease (HD) and Fanconi anemia (FA), a genetic disorder known to cause blood marrow and congenital abnormalities—even cancer.^[footnote]

Link to Learning

Watch this short video to learn more about the founder and bottleneck effects.

https://www.openstax.org/l/founder_bottle

Scientific Method Connection

Testing the Bottleneck Effect

Question: How do natural disasters affect a population's genetic structure?

Background: When an earthquake or hurricane suddenly wipes out much of a population, the surviving individuals are usually a random sampling of the original group. As a result, the population's genetic makeup can change dramatically. We call this phenomenon the bottleneck effect.

Hypothesis: Repeated natural disasters will yield different population genetic structures; therefore,

each time one runs this experiment the results will vary.

Test the hypothesis: Count out the original population using different colored beads. For example, red, blue, and yellow beads might represent red, blue, and yellow individuals. After recording the number of each individual in the original population, place them all in a bottle with a narrow neck that will only allow a few beads out at a time. Then, pour $\frac{1}{3}$ of the bottle's contents into a bowl. This represents the surviving individuals after a natural disaster kills a majority of the population. Count the number of the different colored beads in the bowl, and record it. Then, place all of the beads back in the bottle and repeat the experiment four more times.

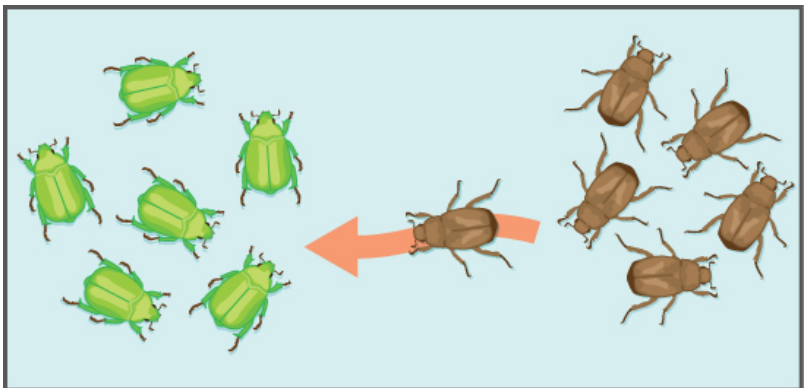
Analyze the data: Compare the five populations that resulted from the experiment. Do the populations all contain the same number of different colored beads, or do they vary? Remember, these populations all came from the same exact parent population.

Form a conclusion: Most likely, the five resulting populations will differ quite dramatically. This is because natural disasters are not selective—they kill and spare individuals at random. Now think about how this might affect a real population. What happens when a hurricane hits the Mississippi Gulf Coast? How do the seabirds that live on the beach fare?

Gene flow can occur when an individual travels from one geographic location to another.

Gene Flow

Another important evolutionary force is **gene flow**: the flow of alleles in and out of a population due to the migration of individuals or gametes ([\[link\]](#)). While some populations are fairly stable, others experience more flux. Many plants, for example, send their pollen far and wide, by wind or by bird, to pollinate other populations of the same species some distance away. Even a population that may initially appear to be stable, such as a pride of lions, can experience its fair share of immigration and emigration as developing males leave their mothers to seek out a new pride with genetically unrelated females. This variable flow of individuals in and out of the group not only changes the population's gene structure, but it can also introduce new genetic variation to populations in different geological locations and habitats.



Mutation

Mutations are changes to an organism's DNA and are an important driver of diversity in populations. Species evolve because of mutations accumulating over time. The appearance of new mutations is the most common way to introduce novel genotypic and phenotypic variance. Some mutations are unfavorable or harmful and are quickly eliminated from the population by natural selection. Others are beneficial and will spread through the population. Whether or not a mutation is beneficial or harmful is determined by whether it helps an organism survive to sexual maturity and reproduce. Some mutations do not do anything and can linger, unaffected by natural selection, in the genome. Some can have a dramatic effect on a gene and the resulting phenotype.

Nonrandom Mating

If individuals nonrandomly mate with their peers, the result can be a changing population. There are many reasons **nonrandom mating** occurs. One reason is simple mate choice. For example, female peahens may prefer peacocks with bigger, brighter tails. Natural selection picks traits that lead to more

mating selections for an individual. One common form of mate choice, called **assortative mating**, is an individual's preference to mate with partners who are phenotypically similar to themselves.

Another cause of nonrandom mating is physical location. This is especially true in large populations spread over vast geographic distances where not all individuals will have equal access to one another. Some might be miles apart through woods or over rough terrain, while others might live immediately nearby.

The temperature at which the eggs are incubated determine the American alligator's (*Alligator mississippiensis*) sex. Eggs incubated at 30°C produce females, and eggs incubated at 33°C produce males. (credit: Steve Hillebrand, USFWS)

Environmental Variance

Genes are not the only players involved in determining population variation. Other factors, such as the environment ([\[link\]](#)) also influence phenotypes. A beachgoer is likely to have darker skin than a city dweller, for example, due to regular exposure to the sun, an environmental factor. For some species, the environment determines some major characteristics, such as gender. For example, some turtles and other reptiles have temperature-dependent sex determination (TSD). TSD means that individuals develop into males if their eggs are

incubated within a certain temperature range, or females at a different temperature range.



Geographic separation between populations can lead to differences in the phenotypic variation between those populations. We see such **geographical variation** between most populations and it can be significant. We can observe one type of geographic variation, a **cline**, as given species' populations vary gradually across an ecological gradient. Species of warm-blooded animals, for example, tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is a latitudinal cline. Alternatively, flowering plants tend to bloom at different times depending on where they are along a mountain slope. This is an altitudinal cline.

If there is gene flow between the populations, the

individuals will likely show gradual differences in phenotype along the cline. Restricted gene flow, alternatively can lead to abrupt differences, even speciation.

Section Summary

Both genetic and environmental factors can cause phenotypic variation in a population. Different alleles can confer different phenotypes, and different environments can also cause individuals to look or act differently. Only those differences encoded in an individual's genes, however, can pass to its offspring and, thus, be a target of natural selection. Natural selection works by selecting for alleles that confer beneficial traits or behaviors, while selecting against those for deleterious qualities. Genetic drift stems from the chance occurrence that some individuals in the gene line have more offspring than others. When individuals leave or join the population, allele frequencies can change as a result of gene flow. Mutations to an individual's DNA may introduce new variation into a population. Allele frequencies can also alter when individuals do not randomly mate with others in the group.

Visual Connection Questions

[\[link\]](#) Do you think genetic drift would happen more quickly on an island or on the mainland?

[\[link\]](#) Genetic drift is likely to occur more rapidly on an island where smaller populations are expected to occur.

Review Questions

When male lions reach sexual maturity, they leave their group in search of a new pride. This can alter the allele frequencies of the population through which of the following mechanisms?

1. natural selection
2. genetic drift
3. gene flow
4. random mating

C

Which of the following evolutionary forces can introduce new genetic variation into a population?

1. natural selection and genetic drift
 2. mutation and gene flow
 3. natural selection and nonrandom mating
 4. mutation and genetic drift
-

B

What is assortative mating?

1. when individuals mate with those who are similar to themselves
 2. when individuals mate with those who are dissimilar to themselves
 3. when individuals mate with those who are the most fit in the population
 4. when individuals mate with those who are least fit in the population
-

A

When closely related individuals mate with each other, or inbreed, the offspring are often not as fit as the offspring of two unrelated individuals. Why?

1. Close relatives are genetically incompatible.
2. The DNA of close relatives reacts

- negatively in the offspring.
3. Inbreeding can bring together rare, deleterious mutations that lead to harmful phenotypes.
 4. Inbreeding causes normally silent alleles to be expressed.
-

C

What is a cline?

1. the slope of a mountain where a population lives
 2. the degree to which a mutation helps an individual survive
 3. the number of individuals in the population
 4. gradual geographic variation across an ecological gradient
-

D

Critical Thinking Questions

Describe a situation in which a population

would undergo the bottleneck effect and explain what impact that would have on the population's gene pool.

A hurricane kills a large percentage of a population of sand-dwelling crustaceans—only a few individuals survive. The alleles carried by those surviving individuals would represent the entire population's gene pool. If those surviving individuals are not representative of the original population, the post-hurricane gene pool will differ from the original gene pool.

Describe natural selection and give an example of natural selection at work in a population.

The theory of natural selection stems from the observation that some individuals in a population survive longer and have more offspring than others: thus, more of their genes are passed to the next generation. For example, a big, powerful male gorilla is much more likely than a smaller, weaker one to become the population's silverback: the pack's leader who mates far more than the other males of the group. Therefore, the pack leader will father more offspring who share half of his genes and are likely to grow bigger and stronger like their father. Over time, the genes for bigger size will

increase in frequency in the population, and the average body size, as a result, will grow larger on average.

Explain what a cline is and provide examples.

A cline is a type of geographic variation that is seen in populations of a given species that vary gradually across an ecological gradient. For example, warm-blooded animals tend to have larger bodies in the cooler climates closer to the earth's poles, allowing them to better conserve heat. This is considered a latitudinal cline. Flowering plants tend to bloom at different times depending on where they are along the slope of a mountain. This is known as an altitudinal cline.

Glossary

assortative mating

when individuals tend to mate with those who are phenotypically similar to themselves

bottleneck effect

magnification of genetic drift as a result of natural events or catastrophes

cline

gradual geographic variation across an ecological gradient

gene flow

flow of alleles in and out of a population due to the individual or gamete migration

genetic drift

effect of chance on a population's gene pool

genetic variance

diversity of alleles and genotypes in a population

geographical variation

differences in the phenotypic variation between populations that are separated geographically

heritability

fraction of population variation that can be attributed to its genetic variance

inbreeding

mating of closely related individuals

inbreeding depression

increase in abnormalities and disease in inbreeding populations

nonrandom mating

changes in a population's gene pool due to mate choice or other forces that cause

individuals to mate with certain phenotypes
more than others

population variation

distribution of phenotypes in a population

selective pressure

environmental factor that causes one
phenotype to be better than another

Adaptive Evolution

By the end of this section, you will be able to do the following:

- Explain the different ways natural selection can shape populations
- Describe how these different forces can lead to different outcomes in terms of the population variation

Natural selection only acts on the population's heritable traits: selecting for beneficial alleles and thus increasing their frequency in the population, while selecting against deleterious alleles and thereby decreasing their frequency. Scientists call this process **adaptive evolution**. Natural selection does not act on individual alleles, but on entire organisms. An individual may carry a very beneficial genotype with a resulting phenotype that, for example, increases the ability to reproduce (fecundity), but if that same individual also carries an allele that results in a fatal childhood disease, that fecundity phenotype will not pass to the next generation because the individual will not live to reach reproductive age. Natural selection acts at the individual's level. It selects for individuals with greater contributions to the gene pool of the next generation. Scientists call this an organism's **evolutionary (Darwinian) fitness**.

Fitness is often quantifiable and is measured by

scientists in the field. However, it is not an individual's absolute fitness that counts, but rather how it compares to the other organisms in the population. Scientists call this concept **relative fitness**, which allows researchers to determine which individuals are contributing additional offspring to the next generation, and thus, how the population might evolve.

There are several ways selection can affect population variation: stabilizing selection, directional selection, diversifying selection, frequency-dependent selection, and sexual selection. As natural selection influences the allele frequencies in a population, individuals can either become more or less genetically similar and the phenotypes can become more similar or more disparate.

Stabilizing Selection

If natural selection favors an average phenotype, selecting against extreme variation, the population will undergo **stabilizing selection** ([\[link\]](#)). In a mouse population that live in the woods, for example, natural selection is likely to favor mice that best blend in with the forest floor and are less likely for predators to spot. Assuming the ground is a fairly consistent shade of brown, those mice whose fur is most closely matched to that color will be most likely to survive and reproduce, passing on

their genes for their brown coat. Mice that carry alleles that make them a bit lighter or a bit darker will stand out against the ground and be more likely to fall victim to predation. As a result of this selection, the population's genetic variance will decrease.

Directional Selection

When the environment changes, populations will often undergo **directional selection** ([\[link\]](#)), which selects for phenotypes at one end of the spectrum of existing variation. A classic example of this type of selection is the evolution of the peppered moth in eighteenth- and nineteenth-century England. Prior to the Industrial Revolution, the moths were predominately light in color, which allowed them to blend in with the light-colored trees and lichens in their environment. However, as soot began spewing from factories, the trees darkened, and the light-colored moths became easier for predatory birds to spot. Over time, the frequency of the moth's melanic form increased because they had a higher survival rate in habitats affected by air pollution because their darker coloration blended with the sooty trees. Similarly, the hypothetical mouse population may evolve to take on a different coloration if something were to cause the forest floor where they live to change color. The result of this type of selection is a shift in the population's genetic variance toward the

new, fit phenotype.

Link to Learning

In science, we sometimes believe some things are true, and then new information becomes available that changes our understanding. The peppered moth story is an example: some scientists recently have questioned the facts behind the selection toward darker moths. Read this [article](#) to learn more.

Diversifying Selection

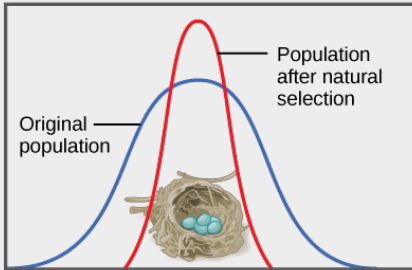
Sometimes two or more distinct phenotypes can each have their advantages for natural selection, while the intermediate phenotypes are, on average, less fit. Scientists call this **diversifying selection** ([\[link\]](#)) We see this in many animal populations that have multiple male forms. Large, dominant alpha males use brute force to obtain mates, while small males can sneak in for furtive copulations with the females in an alpha male's territory. In this case, both the alpha males and the "sneaking" males will be selected for, but medium-sized males, who can't overtake the alpha males and are too big to sneak

populations, are selected against. Diversifying selection can also occur when environmental changes favor individuals on either end of the phenotypic spectrum. Imagine a mouse population living at the beach where there is light-colored sand interspersed with patches of tall grass. In this scenario, light-colored mice that blend in with the sand would be favored, as well as dark-colored mice that can hide in the grass. Medium-colored mice, alternatively would not blend in with either the grass or the sand, and thus predators would most likely eat them. The result of this type of selection is increased genetic variance as the population becomes more diverse.

Visual Connection

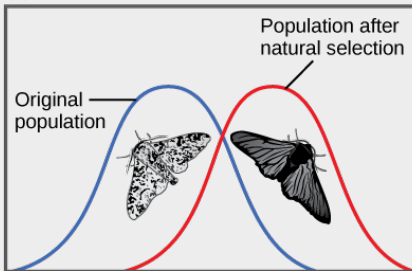
Different types of natural selection can impact the distribution of phenotypes within a population. In (a) stabilizing selection, an average phenotype is favored. In (b) directional selection, a change in the environment shifts the spectrum of observed phenotypes. In (c) diversifying selection, two or more extreme phenotypes are selected for, while the average phenotype is selected against.

(a) **Stabilizing selection**



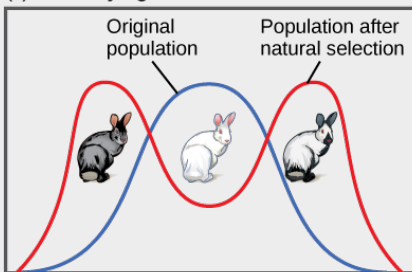
Robins typically lay four eggs, an example of stabilizing selection. Larger clutches may result in malnourished chicks, while smaller clutches may result in no viable offspring.

(b) **Directional selection**



Light-colored peppered moths are better camouflaged against a pristine environment; likewise, dark-colored peppered moths are better camouflaged against a sooty environment. Thus, as the Industrial Revolution progressed in nineteenth-century England, the color of the moth population shifted from light to dark, an example of directional selection.

(c) **Diversifying selection**



In a hypothetical population, gray and Himalayan (gray and white) rabbits are better able to blend with a rocky environment than white rabbits, resulting in diversifying selection.

In recent years, factories have become cleaner, and release less soot into the environment. What impact do you think this has had on the distribution of moth color in the population?

A yellow-throated side-blotched lizard is smaller than either the blue-throated or orange-throated males and appears a bit like the females of the

species, allowing it to sneak copulations. (credit: “tinyfroglet”/Flickr)

Frequency-Dependent Selection

Another type of selection, **frequency-dependent selection**, favors phenotypes that are either common (positive frequency-dependent selection) or rare (negative frequency-dependent selection). We can observe an interesting example of this type of selection in a unique group of Pacific Northwest lizards. Male common side-blotched lizards come in three throat-color patterns: orange, blue, and yellow. Each of these forms has a different reproductive strategy: orange males are the strongest and can fight other males for access to their females. Blue males are medium-sized and form strong pair bonds with their mates. Yellow males ([\[link\]](#)) are the smallest, and look a bit like females, which allows them to sneak copulations. Like a game of rock-paper-scissors, orange beats blue, blue beats yellow, and yellow beats orange in the competition for females. That is, the big, strong orange males can fight off the blue males to mate with the blue’s pair-bonded females, the blue males are successful at guarding their mates against yellow sneaker males, and the yellow males can sneak copulations from the potential mates of the large, polygynous orange males.



In this scenario, natural selection favors orange males when blue males dominate the population. Blue males will thrive when the population is mostly yellow males, and yellow males will be selected for when orange males are the most populous. As a result, populations of side-blotched lizards cycle in the distribution of these phenotypes—in one generation, orange might predominate, and then yellow males will begin to rise in frequency. Once yellow males comprise a majority of the population, blue males will be selected. Finally, when blue males become common, orange males once again

will be favored.

Negative frequency-dependent selection serves to increase the population's genetic variance by selecting for rare phenotypes; whereas, positive frequency-dependent selection usually decreases genetic variance by selecting for common phenotypes.

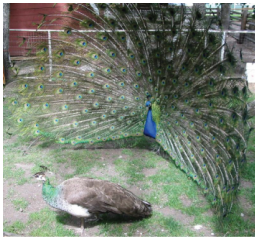
Sexual dimorphism in (a) peacocks and peahens, (b) *Argiope appensa* spiders (the female spider is the large one), and in (c) wood ducks. (credit “spiders”: modification of work by “Sanba38”/Wikimedia Commons; credit “duck”: modification of work by Kevin Cole)

Sexual Selection

Males and females of certain species are often quite different from one another in ways beyond the reproductive organs. Males are often larger, for example, and display many elaborate colors and adornments, like the peacock's tail, while females tend to be smaller and duller in decoration. We call such differences **sexual dimorphisms** ([\[link\]](#)), which arise in many populations, particularly animal populations, where there is more variance in the male's reproductive success than that of the females. That is, some males—often the bigger, stronger, or more decorated males—obtain the vast majority of the total matings, while others receive none. This can occur because the males are better at

fighting off other males, or because females will choose to mate with the bigger or more decorated males. In either case, this variation in reproductive success generates a strong selection pressure among males to obtain those matings, resulting in the evolution of bigger body size and elaborate ornaments to attract the females' attention. Females, however, tend to achieve a handful of selected matings; therefore, they are more likely to select more desirable males.

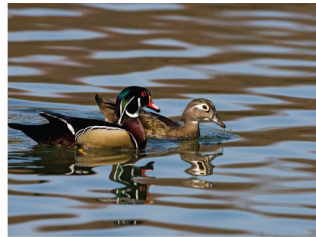
Sexual dimorphism varies widely among species, and some species are even sex-role reversed. In such cases, females tend to have a greater variance in their reproductive success than males and are correspondingly selected for the bigger body size and elaborate traits usually characteristic of males.



(a)



(b)



(c)

We call the selection pressures on males and females to obtain matings sexual selection. It can result in developing secondary sexual characteristics that do not benefit the individual's likelihood of survival but help to maximize its reproductive success. Sexual selection can be so strong that it selects traits that are actually detrimental to the individual's survival. Think, once again, about the peacock's tail.

While it is beautiful and the male with the largest, most colorful tail is more likely to win the female, it is not the most practical appendage. In addition to greater visibility to predators, it makes the males slower in their attempted escapes. There is some evidence that this risk is why females like the big tails in the first place. The speculation is that large tails carry risk, and only the best males survive that risk: the bigger the tail, the more fit the male. We call this the **handicap principle**.

The **good genes hypothesis** states that males develop these impressive ornaments to show off their efficient metabolism or their ability to fight disease. Females then choose males with the most impressive traits because it signals their genetic superiority, which they will then pass on to their offspring. Although one may argue that females should not be picky because it will likely reduce their number of offspring, if better males father more fit offspring, it may be beneficial. Fewer, healthier offspring may increase the chances of survival more than many, weaker offspring.

Link to Learning

In 1915, biologist Ronald Fisher proposed another model of sexual selection: the **Fisherian runaway model**, which suggests that selection of certain traits is a result of sexual preference.

In both the handicap principle and the good genes hypothesis, the trait is an **honest signal** of the males' quality, thus giving females a way to find the fittest mates— males that will pass the best genes to their offspring.

No Perfect Organism

Natural selection is a driving force in evolution and can generate populations that are better adapted to survive and successfully reproduce in their environments. However, natural selection cannot produce the perfect organism. Natural selection can only select on existing variation in the population. It does not create anything from scratch. Thus, it is limited by a population's existing genetic variance and whatever new alleles arise through mutation and gene flow.

Natural selection is also limited because it works at the individual, not allele level, and some alleles are linked due to their physical proximity in the genome, making them more likely to pass on together (linkage disequilibrium). Any given individual may carry some beneficial and some unfavorable alleles. It is the alleles' net effect, or the organism's fitness, upon which natural selection can act. As a result, good alleles can be lost if individuals who carry them also have several overwhelmingly bad alleles. Likewise, bad alleles

can be kept if individuals who have enough good alleles to result in an overall fitness benefit carry them.

Furthermore, natural selection can be constrained by the relationships between different polymorphisms. One morph may confer a higher fitness than another, but may not increase in frequency because going from the less beneficial to the more beneficial trait would require going through a less beneficial phenotype. Think back to the mice that live at the beach. Some are light-colored and blend in with the sand, while others are dark and blend in with the patches of grass. The dark-colored mice may be, overall, more fit than the light-colored mice, and at first glance, one might expect the light-colored mice to be selected for a darker coloration. However, remember that the intermediate phenotype, a medium-colored coat, is very bad for the mice—they cannot blend in with either the sand or the grass and predators are more likely to eat them. As a result, the light-colored mice would not be selected for a dark coloration because those individuals who began moving in that direction (began selection for a darker coat) would be less fit than those that stayed light.

Finally, it is important to understand that not all evolution is adaptive. While natural selection selects the fittest individuals and often results in a more fit population overall, other forces of evolution,

including genetic drift and gene flow, often do the opposite: introducing deleterious alleles to the population's gene pool. Evolution has no purpose—it is not changing a population into a preconceived ideal. It is simply the sum of the various forces that we have described in this chapter and how they influence the population's genetic and phenotypic variance.

Section Summary

Because natural selection acts to increase the frequency of beneficial alleles and traits while decreasing the frequency of deleterious qualities, it is adaptive evolution. Natural selection acts at the individual level, selecting for those that have a higher overall fitness compared to the rest of the population. If the fit phenotypes are those that are similar, natural selection will result in stabilizing selection, and an overall decrease in the population's variation. Directional selection works to shift a population's variance toward a new, fit phenotype, as environmental conditions change. In contrast, diversifying selection results in increased genetic variance by selecting for two or more distinct phenotypes.

Other types of selection include frequency-dependent selection, in which individuals with either common (positive frequency-dependent

selection) or rare (negative frequency-dependent selection) are selected. Finally, sexual selection results from one sex having more variance in the reproductive success than the other. As a result, males and females experience different selective pressures, which can often lead to the evolution of phenotypic differences, or sexual dimorphisms, between the two.

Visual Connection

[\[link\]](#) In recent years, factories have become cleaner, and less soot is released into the environment. What impact do you think this has had on the distribution of moth color in the population?

[\[link\]](#) Moths have shifted to a lighter color.

Review Questions

Which type of selection results in greater genetic variance in a population?

1. stabilizing selection
 2. directional selection
 3. diversifying selection
 4. positive frequency-dependent selection
-

C

When males and females of a population look or act differently, it is referred to as _____.

1. sexual dimorphism
 2. sexual selection
 3. diversifying selection
 4. a cline
-

A

The good genes hypothesis is a theory that explains what?

1. why more fit individuals are more likely to have more offspring
2. why alleles that confer beneficial traits or behaviors are selected for by natural selection
3. why some deleterious mutations are maintained in the population
4. why individuals of one sex develop

D

Critical Thinking Questions

Give an example of a trait that may have evolved as a result of the handicap principle and explain your reasoning.

The peacock's tail is a good example of the handicap principle. The tail, which makes the males more visible to predators and less able to escape, is clearly a disadvantage to the bird's survival. But because it is a disadvantage, only the most fit males should be able to survive with it. Thus, the tail serves as an honest signal of quality to the females of the population; therefore, the male will earn more matings and greater reproductive success.

List the ways in which evolution can affect population variation and describe how they influence allele frequencies.

There are several ways evolution can affect population variation: stabilizing selection, directional selection, diversifying selection, frequency-dependent selection, and sexual selection. As these influence the allele frequencies in a population, individuals can either become more or less related, and the phenotypes displayed can become more similar or more disparate.

Glossary

adaptive evolution

increase in frequency of beneficial alleles and decrease in deleterious alleles due to selection

directional selection

selection that favors phenotypes at one end of the spectrum of existing variation

diversifying selection

selection that favors two or more distinct phenotypes

evolutionary fitness

(also, Darwinian fitness) individual's ability to survive and reproduce

frequency-dependent selection

selection that favors phenotypes that are either common (positive frequency-dependent

selection) or rare (negative frequency-dependent selection)

good genes hypothesis

theory of sexual selection that argues individuals develop impressive ornaments to show off their efficient metabolism or ability to fight disease

handicap principle

theory of sexual selection that argues only the fittest individuals can afford costly traits

honest signal

trait that gives a truthful impression of an individual's fitness

relative fitness

individual's ability to survive and reproduce relative to the rest of the population

sexual dimorphism

phenotypic difference between a population's males and females

stabilizing selection

selection that favors average phenotypes

The Periodic Table of Elements

Periodic Table of the Elements

Group

12

1

H

1.008

hydrogen

2

He

4.003

helium

3

Li

6.94

lithium

4

Be

9.012

beryllium

11

Na

22.99

sodium

12

Mg

24.31

magnesium

13

B

10.81

boron

14

C

12.01

carbon

15

N

14.01

nitrogen

16

O

16.00

oxygen

17

F

19.00

fluorine

18

Ne

20.18

neon

19

K

39.10

potassium

20

Ca

40.08

calcium

21

Sc

44.96

scandium

22

Ti

47.87

titanium

23

V

50.94

vanadium

24

Cr

52.00

chromium

25

Mn

54.94

manganese

26

Fe

55.85

iron

27

Co

58.93

cobalt

28

Ni

58.69

nickel

29

Cu

63.55

copper

30

Zn

65.38

zinc

31

Ga

69.72

gallium

32

Ge

72.63

germanium

33

As

74.92

arsenic

34

Se

78.97

selenium

35

Br

79.90

bromine

36

Kr

83.80

krypton

37

Rb

85.47

rubidium

38

Sr

87.62

strontium

39

Y

88.91

yttrium

40

Zr

91.22

zirconium

41

Nb

92.91

niobium

42

Mo

95.95

molybdenum

43

Tc

98.91

technetium

44

Ru

101.1

ruthenium

45

Rh

102.9

rhodium

46

Pd

106.4

palladium

47

Ag

107.9

silver

48

Cd

112.4

cadmium

49

In

114.8

indium

50

Sn

118.7

tin

51

Sb

121.8

antimony

52

Te

127.6

tellurium

53

I

126.9

iodine

54

Xe

131.3

xenon

55

Cs

132.9

cesium

56

Ba

137.3

barium

57-71

La-Lu

lanthanum series

72

Hf

168.9

hafnium

73

Ta

180.9

tantalum

74

W

183.8

tungsten

75

Re

186.2

rhenium

76

Os

190.2

osmium

77

Ir

192.2

iridium

78

Pt

195.1

platinum

79

Au

197.0

gold

80

Hg

200.6

mercury

81

Tl

204.4

thallium

82

Pb

207.2

lead

83

Bi

209.0

bismuth

84

Po

[209]

polonium

85

At

[210]

astatine

86

Rn

[222]

radon

87

Fr

[223]

francium

88

Ra

[226]

radium

89-103

Ac-Lr

actinium series

104

Rf

[267]

rutherfordium

105

Sg

[271]

seaborgium

106

Bh

[273]

bohrium

107

Hs

[277]

hassium

108

Mc

[288]

moscovium

109

Mt

[276]

meitnerium

110

Ds

[281]

darmstadtium

111

Rg

[285]

roentgenium

112

Cn

[285]

copernicium

113

Nh

[289]

nihonium

114

Fl

[289]

flerovium

115

Mc

[289]

moscovium

116

Lr

[293]

livermorium

117

Ts

[294]

tennessine

118

Og

[294]

oganesson

57

La

138.9

lanthanum

58

Ce

140.1

cerium

59

Pr

140.9

praseodymium

60

Nd

144.2

neodymium

61

Pm

[145]

promethium

62

Sm

150.4

samarium

63

Eu

151.9

europium

64

Gd

157.3

gadolinium

65

Tb

158.9

terbium

66

Dy

162.5

dysprosium

67

Ho

164.9

holmium

68

Er

167.3

erbium

69

Tm

168.9

thulium

70

Yb

173.0

ytterbium

71

Lu

174.9

lutetium

89

Fr

[223]

francium

90

Th

232.0

thorium

91

Pa

231.0

protactinium

92

U

238.0

uranium

93

Np

[237]

neptunium

94

Pu

[244]

plutonium

95

Am

[243]

americium

96

Cm

[247]

curium

97

Bk

[247]

berkelium

98

Cf

[251]

californium

99

Es

[252]

einsteinium

100

Fm

[257]

fermium

101

Md

[258]

mendelevium

102

No

[259]

nobelium

103

Lr

[262]

lawrencium

Atomic number

1

Symbol

H

Atomic mass

1.008

Name

hydrogen

Color Code

Metal

Metalloid

Nonmetal

Solid

Liquid

Gas

Measurements and the Metric System

Measurements and the Metric System

Measurement	Unit	Abbreviation	Metric Equivalent	Approximate Standard Equivalent
Length	1 nanometer	nm	1 nm = 10^{-9} m	
	1 centimeter	cm		
	1 cm = 0.394 inch			
	1 m = 39.37 inches			
	1 m = 3.28 feet			
	1 m = 1.093 yards			
	1 km = 0.621 miles			
micrometer	μ m	1 μ m = 10^{-6} m		
millimeter	mm	1 mm = 0.001 m		
centimeter	cm	1 cm = 0.01 m		
meter	1 mm = 100 cm			
	1 m = 1000 mm			
kilometer	km	1 km = 1000 m		
Mass	1 gram	g	1 μ g =	
	1 microgram	μ g		

		1 kg = 2.205 pounds	10^{-6} g
milligram	mg	1 mg =	
		10^{-3} g	
gram	g	1 g =	
		1000 mg	
kilogram	kg	1 kg =	
		1000 g	
Volume		1 milliliter = 0.034 fluid ounce	1 μ l =
		1 l = 1.057 quarts	10^{-6} l
		1 kl = 264.172 gallons	
milliliter	ml	1 ml =	
		10^{-3} l	
liter	l	1 l =	
		1000 ml	
kiloliter	kl	1 kl =	
		1000 l	
Area		1 square centimeter = 0.155 square inch	1 cm ² =
		1 centimeter = 0.394 square feet	100 mm ²
		1 m ² = 1.196 square yards	
		1 ha = 2.471 acres	
square meter	m ²	1 m ² =	
		10,000	
		cm ²	
hectare	ha	1 ha =	
		10,000 m ²	
Temperature	Celsius	°C	—
			1 °C =
			$5/9 \times (°F$
			$- 32)$